PB85-187680

Survey Management Handbook. Volume 2 Overseeing the Technical Progress of a Survey Contract

(U.S.) Environmental Protection Agency
Washington, DC

Dec 84

U.S. Department of Commerce
National Technical Information Service
NTIS

*when published

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-230/12-84-002	2.	3. RECIPIENT'S ACCESSION NO. PB85 187680/AS
4. TITLE AND SUBTITLE SURVEY MANAGEMENT HANDBOOK: Volume I: Guidelines for Planning and Managing a Statistical Survey SURVEY MANAGEMENT HANDBOOK: Volume II: Overseeing the Technical Progress of a Survey Contract		5. REPORT DATE Vol. I Nov. 83* Vol. II Dec. 84*
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Cynthia Croce		8. PERFORMING ORGANIZATION REPORT NO. N/A
9. PERFORMING ORGANIZATION NAME AND ADDRESS Environmental Protection Agency Statistical Policy Branch (PM-223) 401 M Street, SW Washington, DC 20460		10. PROGRAM ELEMENT NO. N/A
		11. CONTRACT/GRANT NO. N/A
12. SPONSORING AGENCY NAME AND ADDRESS Same as #9		13. TYPE OF REPORT AND PERIOD COVERED Handbook
		14. SPONSORING AGENCY CODE EPA 400/00

15. SUPPLEMENTARY NOTES

16. ABSTRACT

Volume I focuses on survey design principles and ways program officials might productively apply them in planning a contract survey.

Volume II emphasizes the conduct and management of an Agency-sponsored survey. Each of th six chapters of Volume II corresponds to a major component of a typical work plan for a statistical survey of human populations.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
statistical surveys, human populations, questionnaires, sampling plan, interviewing		

18. DISTRIBUTION STATEMENT Release Unlimited	19. SECURITY CLASS *(This Report)* UNCLAS	21. NO. OF PAGES Vol. I - 145 Vol. II - 175
	20. SECURITY CLASS *(This page)* UNCLAS	22. PRICE

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

I

United States
Environmental Protection
Agency

Office of Policy,
Planning, and Evaluation
Washington, DC 20460

EPA-230/12-84-002
December 1984

PB85-187680

EPA

Survey Management Handbook

Volume II: Overseeing the Technical Progress of a Survey Contract

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

For additional copies, please contact:

N. PHILLIP ROSS, Chief,
Statistical Policy Branch,
Office of Standards and Regulations
U.S. Environmental Protection Agency
PM-223, 401 M Street, S.W.,
Washington, D.C. 20460

December 1984

SURVEY MANAGEMENT HANDBOOK STAFF

Project Manager	MEL KOLLANDER, EPA
Principal Writer	CYNTHIA CROCE, Consultant
Statistical Advisor	THOMAS B. JABINE, Consultant, Committee on National Statistics, National Academy of Science
Editor and Proofreader ..	PATRICIA MINAMI, EPA

ENVIRONMENTAL PROTECTION AGENCY

VOLUME II

SURVEY MANAGEMENT HANDBOOK

Overseeing the Technical Progress of a Survey Contract

TABLE OF CONTENTS

Page

TABLE OF CONTENTS i

TABLE OF EXHIBITS vii

INTRODUCTION 1

CHAPTER 1 - FROM DESIGN TO ANALYSIS 5

A - APPROACHES USED TO ANALYZE SURVEY DATA 5

B - STEPS IN PREPARING AN ANALYSIS PLAN 7

1. Define the Purpose of the Survey 9
2. Define the Research Objectives 11
3. Define the Study Variables 13
4. Specify the Analytic Approaches and Methods . 14
5. Define the Preliminary Tabulations 15

CHAPTER 2 - SELECTING THE DATA COLLECTION METHOD 19

A - PRINCIPAL DATA COLLECTION METHODS 20

1. Traditional Survey Research Methods 20
2. Exploratory Research Methods 22

B - COMPARISON OF THE THREE TRADITIONAL COLLECTION METHODS 24

1. Special Characteristics of Face-to-Face Surveys 24
2. Special Characteristics of Telephone Surveys 26
3. Special Characteristics of Mail Surveys . . . 29

C - FACTORS AFFECTING THE CHOICE OF COLLECTION METHODS 30

1. Characteristics of the Target Population . . 31
2. Data Requirements 31
3. Respondent's Obligation to Reply 32
4. Target Response Rate 32
5. Available Time 32
6. Available Funds 33

D - ASSESSING THE SUITABILITY OF THE PROPOSED COLLECTION METHOD 33

Page

CHAPTER 3 - DEVELOPING THE QUESTIONNAIRE 37

A - STEPS IN THE DEVELOPMENT OF A SURVEY QUESTIONNAIRE 37

1. Determine the Analysis Requirements 40
2. Draft a List of Topics or Suggested Questions 40
3. Conduct Exploratory Interviews with a Few Individuals in the Population 41
4. Prepare First Draft of the Questionnaire . . 42
5. Review and Approve First Draft of the Questionnaire 43
6. Prepare Plan for Pretest 44
7. Initiate Clearance Request for the Pretest . 46
8. Pretest on a Sample of the Target Population 46
9. Debrief Interviewers and Assess Pretest Findings 48
10. Revise Questionnaire and Prepare Plan for the Pilot Test 48
11. Review Revised Questionnaire and Pilot Test Plan 50
12. Recruit Interviewers and Prepare Training Materials 50
13. Pilot Test Questionnaire and Assess Results . 51
14. Revise Questionnaire and Collection Procedures for Main Survey 51
15. Review and Approve Procedures for the Main Survey 52
16. Print Questionnaire 52

B - REVIEWING DRAFT QUESTIONNAIRES 53

1. Reviewing Individual Questions 53
2. Reviewing the Overall Content and Organization 64
3. Reviewing the Format 69

C - MONITORING PRETESTS 73

CHAPTER 4 - SAMPLING 77

A - ADVANTAGES OF USING SAMPLING 77

1. Lower Costs 78
2. Reduced Paperwork Demands 78
3. More Timely Results 79
4. More Accurate Results 79

Preceding page blank

Page

CHAPTER 4 - SAMPLING (Continued)

B - SAMPLING ERRORS AND SAMPLE SIZE 80

1. Sampling Errors 80
2. Measuring and Expressing Sampling Errors . . 81
3. Determining Sample Size 82

C - SAMPLING METHODS 85

1. Probability Sampling Methods 86
2. Non-Probability Sampling Methods 95

D - MAJOR COMPONENTS OF A SAMPLING PLAN 98

1. Sampling Frames 98
2. Sample Selection Procedures 100
3. Estimation Procedures 101
4. Calculation of Sampling Errors 105

E - MONITORING THE SAMPLING ACTIVITIES 106

CHAPTER 5 - INTERVIEWING 111

A - ESTABLISHING THE QUALITY-ASSURANCE PROCEDURES . . 111

1. Respondent Rules 112
2. Follow-Up Rules 113
3. Quality Control 114

B - STAFFING AND ORGANIZING THE FIELD OPERATIONS . . 123

1. Preparing Instructions and Training Materials 123
2. Staffing the Field Operations 126
3. Training the Interviewers 129
4. Coordinating and Controlling the Fieldwork 130

C - CONDUCTING THE INTERVIEWS 131

1. Locating Respondents 132
2. Gaining Responsents' Cooperation. 132
3. Asking Questions 134
4. Recording and Editing Responses 135

D - MONITORING THE INTERVIEW PROCESS 135

Page

CHAPTER 6 - DATA PROCESSING 139

A - STEPS IN PROCESSING SURVEY DATA 139

1. Develop the Processing Procedures 140
2. Select and Train Staff 141
3. Screen the Questionnaires 143
4. Review and Edit the Questionnaires. 143
5. Code Open Questions 144
6. Enter Data 145
7. Detect and Resolve Errors in the Data File. . 146
8. Prepare the Outputs 150

B - MONITORING THE DATA PROCESSING ACTIVITIES 155

GLOSSARY 159

LIST OF RECOMMENDED SOURCES 167

TABLE OF EXHIBITS

No. Title Page

- Components of the Work Plan 3

1 Guide for Choosing a Data Collection Method 34

2 The Sponsoring Office's Tasks in the Questionnaire-Development Process 39

3 Criteria for Reviewing Survey Questionnaires 54

4 Absolute and Relative Sampling Errors for Different Types of Estimates of Families Using Contaminated Drinking Water Sources 83

5 Multi-State Design for a National Household Survey . 94

Preceding page blank

INTRODUCTION

Statistical surveys are playing an increasingly important role in Agency decisionmaking. As policymakers demand more quantitative support for Agency decisions, program managers are giving careful consideration to statistical survey reports and their implications in the framing of regulatory decisions and long-range environmental policies. Reliable survey data on the duration, magnitude, and physical distribution of pollutants in the environment have proven invaluable for determining the precise degree of pollutant control needed to respond to various statutory mandates and the manner in which the Agency should exercise such control.

There have been extraordinary advances in survey methodology in the past two decades, the most striking of them in sampling, data processing, and statistical analysis. This has made large-scale collections of demographic and economic facts easier, faster, less costly, and more reliable. Moreover, the quality of reporting both survey methods and survey results has consistently improved. These advances have motivated those who sponsor surveys to demand increasingly higher standards in questionnaire design, data collection methodology, sampling, interviewing, data processing, and analysis.

The growing reliance on high-quality statistical work for Agency planning and policymaking, coupled with the recent advances in survey methodology, in fact, prompted the development of this two-volume Survey Management Handbook.

In Volume I of the Handbook, published in November of 1983, we focused on survey design principles and ways program officials might productively apply them in planning a contract survey. In the present volume, our emphasis is on the conduct and management of an Agency-sponsored survey. Specifically, we examine --

- The methods, procedures, and quality-assurance techniques typically used to collect, process, and analyze survey data; and
- The actions EPA project officials can take to ensure the technical soundness of all contract work performed during the course of a survey.

Volume II is organized into six chapters, which correspond to the major components of a typical work plan for a statistical survey of human populations. Normally the work plan -- and the subsequent fieldwork, data processing, and often the analysis -- is done by a large survey research contractor, with the EPA sponsoring office playing an oversight role throughout the term of the contract.

The work plan establishes the methods and procedures to be used in collecting, processing, and analyzing data from or about the survey population. Usually it consists of --

- An analysis plan
- Specification of the method(s) of collection
- A draft questionnaire and specifications for any pretests
- A sampling plan
- Interviewing procedures
- Data processing procedures

COMPONENTS OF THE WORK PLAN

A summary of the topics covered in each of the six chapters is given on the next page.

As in the previous volume, the survey methods and techniques we discuss are applicable to fairly large-scale surveys. This is because most of EPA's demographic, economic, and social investigations as well as field studies deal with large populations and issues that the Agency must necessarily view from a national perspective.

Of course, not every empirical research project EPA undertakes requires the formal apparatus needed for a large-scale survey. Sometimes it is more appropriate to study a handful of cases intensively rather than investigate a representative sample, to interview a few individuals or groups informally rather than use the structured interviews prescribed for major statistical surveys, or to develop in-depth descriptions of a few individuals rather than aim for a set of statistics about a group. In fact, several approaches may be used to resolve a particular survey research problem. The researcher's challenge is to identify those approaches that are most likely to achieve the specific objectives of the project. The purpose of this Handbook is to help you meet this challenge.

Throughout, we discuss theoretical issues in very general terms. No background knowledge of statistics is presumed. In the event you wish to delve further into survey theory, a list of excellent sources is given at the end of each chapter. A complete

COMPONENTS OF THE WORK PLAN

ANALYSIS PLAN	Chapter 1 examines the steps involved in defining the research objectives of the survey and choosing the analytic approach most appropriate for achieving these objectives.
DATA COLLECTION METHOD	Chapter 2 describes the principal methods of collecting survey data and the factors influencing the choice of methods, and suggests ways of evaluating the method proposed for a particular EPA survey.
QUESTIONNAIRE AND PRETESTING PROCEDURES	Chapter 3 examines the sequence of steps involved in developing a sound survey questionnaire, presents criteria for reviewing draft questionnaires, and recommends ways of monitoring pretests.
SAMPLING PLAN	Chapter 4 describes the advantages of sampling, the principal methods of choosing a sample, and the components of a sampling plan, and recommends ways of monitoring the sampling activities.
INTERVIEWING PROCEDURES	Chapter 5 discusses the administrative and quality-assurance procedures typically used to organize, manage, and monitor a survey where interviewing is used to collect the majority of the data.
DATA PROCESSING PROCEDURES	Chapter 6 looks at the steps involved in processing the raw data collected from the sample to produce tabulations and analyses that will achieve the research objectives.

list of these sources appears at the end of the Handbook along with a glossary of terms.

We strongly suggest that you have a survey statistician review your survey design and analysis plan early in the planning stage -- certainly before you take steps to procure outside technical support. You also may find it necessary to get the advice of experts at various points of the survey in order to effectively apply the methods and techniques we recommend, especially with respect to sampling and data analysis. All too frequently, statisticians are called in after the data are collected, given a stack of completed questionnaires, and asked to make what they can of them. Unfortunately, because of gaps and omissions in the data, flaws in the survey design, mistakes in the questionnaire, and other problems that could easily have been avoided if a survey expert had been called in during the planning stage, there is very little that can be done.

Keep in mind that the Statistical Policy Branch (SPB) of the Chemicals and Statistical Policy Division within the Office of Standards and Regulations, which prepared this Handbook, offers technical assistance to the programs in all facets of survey management.

CHAPTER 1

FROM DESIGN TO ANALYSIS

In a given research situation, survey designers usually have a choice of research designs, methods of observation, methods of measurement, and types of analysis. All must fit together and be appropriate to the research problem. The choices the researchers make in each case will depend on how much is already known about the problems they are investigating and the specific reasons the information is needed.

Whether, as the survey sponsors, you intend to collect descriptive facts about a population or to delve deeper and attempt to explain certain facts requires a clear understanding of what you expect the research effort to achieve. Collecting data in the field is no substitute for well-thought out decisions beforehand about what is, and what is not, worth investigating. Without a clear idea of the objectives of your research, the survey is likely to result in much wasted time and money and the accumulation of much unwanted data.

In this chapter, we discuss --

- The general approaches survey statisticians use to analyze and interpret survey data;
- How to develop an analysis plan that will clearly define the purpose of your survey, the research objectives, the type of data to be collected, and the most appropriate method of analysis for achieving your research objectives; and
- The major components of the work plan, around which this volume is organized. In a contract survey, the work plan describes the methods and procedures the contractor plans to use to collect, process, and analyze the survey data.

A. APPROACHES USED TO ANALYZE SURVEY DATA

In survey research, analysis means categorizing, ordering, manipulating, and summarizing data to obtain answers to research questions. The purpose of analysis is to reduce data to intelligible and interpretable form. The data first are broken down into constituent parts to obtain answers to research questions and test research hypotheses.

Analyzing data does not provide answers to research questions. Interpretation is necessary. To interpret is to explain. Interpretation takes the results of the data analysis, makes inferences relevant to the relationships among the data, and draws conclusions about these relationships. The researcher who makes the interpretation searches the results for their meaning and implications.

A host of analysis techniques are available for studying survey data. However, here we will focus on the four main approaches to analysis, which are --

- Qualitative analysis and evaluation
- Statistical descriptions
- Statistical inference
- Analytic interpretation

APPROACHES
TO DATA
ANALYSIS

We'll discuss each of these approaches briefly in the order of their complexity and sophistication.

(1) Qualitative analysis and evaluation.

In a qualitative analysis, the researcher's goal is to understand the characteristics of a few individuals, rather than the characteristics of a population or subgroups of that population. A qualitative approach generally is not indicated for sample surveys, which are of major interest in this Handbook, but it may be the most suitable approach in some research situations.

For example, non-quantitative analysis is often the preferred approach (a) for analyzing the results of case studies (or field studies) where a relatively small number of individuals (or specimens) are being investigated; (b) for evaluating the results of informal research prior to conducting a full-scale statistical survey; and (c) for developing hypotheses to test in a pilot study or a full-scale survey.

(2) Statistical descriptions.

Statistical descriptions are by far the most common method of reporting survey data. They often are referred to as "statistical analysis," but this fundamental approach to the analysis of survey data simply involves working out statistical distributions, constructing tables and graphs, and calculating simple measures such as means, medians, measures of dispersion, percentages, proportions, etc. It can be used to describe data collected from a probability sample or an entire study population.

In other words, statistical descriptions are the tabulations researchers prepare after the data are processed to aggregate the features of the data file so the analysts can view the database in some intelligible and interpretable form. Statistical descriptions often are done in series, one variable or research question at a time being cross-classified with others, thus producing a descriptive summary of the relationships between the study variables.

(3) Statistical inference.

In the broadest sense of the word, inference is the principal approach for analyzing statistical data. Inference is brought into play whenever data are collected from a probability sample rather than an entire population. When a probability sample is used, the researchers must estimate the population characteristics from those of the sample as well as estimate sampling errors. Statistical inference is the linking of the results derived from data collected from or about a sample to the population from which the sample was drawn.

(4) Analytic interpretation.

This last and most complex approach is a form of statistical inference called analytic interpretation. It refers to the statistician's attempts to explain the relationships between variables using various statistical analysis techniques. For example, researchers may employ a multivariate regression analysis technique to better understand the relationships between exposure to a particular pollutant and the socio-economic characteristics of a study population.

B. STEPS IN PREPARING AN ANALYSIS PLAN

In this section we will show you how to construct an analysis plan to complement the design specifications you establish for your survey. The basic criteria for the survey design and the analysis plan should be developed simultaneously early in the planning stage. Constructing a well thought-out analysis plan will help you define the design criteria so that you can achieve your research objectives with some desired level of accuracy considering the resources you have available. These design criteria and the analysis plan together provide a sound conceptual framework for whatever work you and the contractor subsequently do during the course of the survey.

In Volume 1, we described the sponsoring office's responsibility for defining the following minimum design criteria for the survey along with clear statements of the purpose and objectives of the research.

- Target population and coverage
- Specific data needs
- Use of probability sampling
- Sampling error (precision)
- Target response rate

MINIMUM
SURVEY DESIGN
SPECIFICATIONS

The intent of these criteria is to guide the project staff in developing the statement of work to procure whatever outside technical support may be necessary and to help the contractor prepare a technically - and statistically - sound work plan. They may possibly be modified during the contract negotiations before being incorporated into the contract. We will not further elaborate on these minimum design specifications because they were amply covered in Chapters 3 and 5 of Volume I.

Constructing the analysis plan is a five-step process. The plan should be developed by the project office with the assistance of Agency statisticians, computer programmers, specialists in the subject area of the research, and systems analysts, as appropriate.

The end-products of the five steps, discussed below, are clear definitions of (1) the purpose of the survey, (2) the objectives of the research (the main areas of investigation), (3) the data or the variables to be investigated, (4) the analytic approaches and methods to be used to achieve the research objectives, and (5) the preliminary tabulations to be prepared from the completed data file after the data are processed.

Later, after the Agency and the contractor have studied the preliminary tabulations, the analysis plan can be refined -- usually this is done by the contractor -- to include specifications for additional, perhaps more sophisticated tabulations and the types of statistical analysis techniques that should be applied to fully reveal the informational content of the data base.

- Step 1: Define the Purpose of the Survey

Generally speaking, to define the purpose of a survey is to give the specific reasons why certain information is needed. For any EPA-sponsored survey the reasons must relate to some specific legislative, regulatory, or judicial mandate that either directs the Agency to explore a particular environmental problem or to take certain corrective actions, and EPA cannot faithfully comply with the mandate unless some new or additional empirical information is collected and analyzed.

From a practical standpoint stating the purpose of the survey means defining its operational usefulness for planning and policy analysis. The statement of purpose in your analysis plan, therefore, must clearly show how the data you plan to collect will result in information that will clarify or resolve some specific environmental problem that some authority has directed EPA to deal with. In other words, you must specify --

- How the information is to be used
- The problems to be addressed
- Their relationship to a specific EPA mandate

PURPOSE
OF THE
SURVEY

Below is a statement of purpose that appeared in a recent report on an EPA field study of carbon monoxide (CO) using hand-held personal exposure monitors to test levels of CO in a variety of commercial settings. The survey was conducted by EPA staff in the Office of Monitoring Systems and Quality Assurance, Office of Research and Development. The statement clearly shows how the study results will be applied for planning and policymaking purposes, the problems the researchers intend to deal with, and their relationship to a specific EPA mandate.

> "The goal of air pollution control programs in the U.S., as mandated by Federal law and implemented by the States, is to attain National Ambient Air Quality Standards (NAAQS). The NAAQS for carbon monoxide (CO), for example, specify two different

concentrations and averaging times, neither of which is to be exceeded more than once per year:

35 parts per million (PPM) for 1 hour
9 ppm for 8 hours.

"Both standards are intended to protect against the accumulation of more than 2% carboxyhemoglobin in the blood....

"Nondispersive infrared (NDIR) monitoring at fixed stations is the usual way for determining a given city's compliance with the NAAQS for CO. During the past decade, a number of studies have revealed that concentrations observed at fixed air monitoring stations have not been representative of concentrations sampled throughout an urban area. Some field studies have shown, for example, that commuters in traffic and pedestrians on downtown streets encountered CO levels above the NAAQS on a given date, while official air monitoring stations reported CO values below the NAAQS at the same time. Furthermore, studies of human activities suggest that most people spend the greatest proportion of any given 24-hour period indoors -- in residences, stores, offices, factories, etc. These settings are not necessarily identical to sites selected for fixed air monitoring stations.

"These studies have raised questions about the usefulness of data generated by today's monitoring stations for protection of public health. An unanswered question is the degree to which conventional fixed stations either underestimate or overestimate the actual exposure of people as they go about their daily activities. The studies have stimulated interest in 'exposure monitoring,' which treats the person as a receptor and measures the pollutant levels actually contacting the person's body....

"Prior to the late 1970's there was no low cost, accurate means available for measuring CO concentrations to which people ordinarily were exposed in their daily lives. The advent of microelectronics has brought considerable progress in developing reliable, compact air quality monitoring instruments that can operate on batteries. The most dramatic of these are the new miniaturized personal exposure monitors (PEM's) .... The present investigation is the first large-scale microenvironmental field study to make use of the new CO PEM instruments...."

Since the kinds of problems EPA has been directed to explore and manage encompass such a wide range of health and environmental issues, you may find it relatively easy to develop an adequate statement of purpose for your survey. What normally is far more difficult is building a set of arguments to justify the expenditure of program funds for your particular project, given the limited resources available to each program to address a mind-boggling number of priority issues. A comprehensive, well-reasoned analysis plan will help you build just such a set of arguments.

- Step 2: Define the Research Objectives

Once you have justified the need for the survey from a planning or policymaking standpoint, you can begin to think about how to define its usefulness in "scientific" terms. The desired result should be a clear statement of the research objectives in terms of the --

- Kinds of questions you want answered
- Hypotheses to be tested
- Information to be collected

RESEARCH OBJECTIVES

Continuing with the previous example, lets look at how the objectives of the PEM CO study were framed. EPA staff defined several sets of research questions.

The first set of research questions addressed the CO concentrations typically found in commercial settings, e.g. --

> "What levels of CO ordinarily are present in typical commercial settings?"

> "Are CO levels in typical commercial settings usually zero, negligible, or above the NAAQS?"

The second set of questions concerned the variability of CO concentrations and factors that may be associated with that variability. Examples from this set of questions are --

> "How do CO concentrations vary over time within and between different cities for a given commercial setting?"

> "If CO is a street-level pollutant associated with vehicular traffic, do workers have greater protection in offices on the upper floors of a high-rise building?"

Another set of research questions addressed the accuracy of the fixed-station monitors operated by air quality management districts to measure the air pollution to which the public is actually exposed, e.g. --

> "Do CO concentrations measured in commercial settings using PEM's correlate with ambient concentrations measured at fixed stations using NDIR instruments?"

There also was a set of questions concerning the research methodology itself, including the following items --

> "Is the CO PEM an effective tool for sampling air quality at a variety of urban locations?"
>
> "What are the implications of the present study for future research on exposures of the population to CO?"

Several hypotheses were formed and tested. For example, the researchers tested to see if the indoor concentrations were appreciably less than the outdoor concentrations when the entrance door to each commercial setting was closed.

The information to be collected was identified as --

> "5,000 concentrations of CO at one-minute intervals using PEM's for instantaneous measurement in a variety of commercial settings in several California cities over a nine-month period."

Ultimately five principal objectives were framed:

(1) "To determine the CO concentrations typically found in commercial settings";

(2) "To determine the variability of CO concentrations in commercial settings and the time and spatial factors that may be associated with that variability";

(3) "To define and classify microenvironments which are applicable to commercial settings";

(4) "To determine how accurately fixed station monitors measure the CO settings"; and

(5) "To develop research methodology for measuring CO concentrations in field surveys using PEM's."

When you frame your research objectives be sure they are both specific and answerable. For example, a question like "Is water contaminated by aldicarb?" is not answerable. However, the following is a question researchers can undertake to answer: "What proportion of the U.S. population is consuming water that contains more than seven parts per million (ppm) of aldicarb?" This question, in fact, was an attempt to frame the objectives of an EPA-sponsored field study concerning the pesticide aldicarb, which was believed to be contaminating drinking water in certain communities. Later, because of time and budget constraints, it was reframed as follows, "What proportion of the households in high-risk areas are drinking water with more than seven ppm of aldicarb, where high-risk is defined to be either in counties growing crops that are licensed for aldicarb or in which sales are reported."

It is impossible to overestimate the importance of framing the research objectives of your survey fully and precisely. No amount of data manipulation later can overcome the problems that may result from poorly-defined objectives. Furthermore, once you have defined them, do not attempt to broaden their scope with further research topics or include other types of information unless you are sure of achieving your initial objectives with the resources you have available.

• Step 3: Define the Study Variables

Once the objectives are clearly defined, the next step is to define the key variables of the study. In other words, you will have to identify the specific data items that will be required to meet your stated objectives. A variable is a characteristic of a sample or of a population that varies in magnitude. In surveys of human populations, common variables are age, sex, race, income level, education level, etc.

Returning to our CO PEM example, the basic variable was --

> "the average (mean) of two simultaneously taken one-minute samples of CO concentrations."

Other variables were developed to test different hypotheses such as those used for comparing indoor and outdoor CO concentrations using different settings of the personal exposure monitor and door entrances of the commercial establishments open and closed, e.g. --

> "mean CO concentration of indoor PEM setting i with entrance door closed";
>
> "mean CO concentration of outdoor setting i with entrance door closed";
>
> "mean CO concentration of indoor setting j with entrance door open"; and
>
> "mean CO concentration of outdoor setting j with entrance door open."

- Step 4: Specify the Analytic Approaches and Methods

Following the guidelines we provided in section A, the next step in developing the analysis plan is to determine which analytic approach will allow you to achieve your research objectives most efficiently given the time and resources you have available. This means determining which analysis methods are most likely to achieve each of your research objectives. Note that different observation methods, measurement techniques, and analysis methods may be needed to fulfill each one of your research objectives.

For most studies of human populations, a questionnaire is the basic information gathering tool. If you choose this "method of observation," you may want to prepare a list of preliminary questions that will measure the magnitude of the study variables you identified in the previous step (see Chapter 3 for details on preparing a questionnaire). You'll also have to decide what level of accuracy (or precision) you will require. As discussed in Chapter 3 of Volume I, the level of accuracy you determine should depend on how you plan to use the results of the survey. And, finally, you'll have to determine what minimally acceptable rate of response (target response rate) is necessary to achieve your research objectives. (See Chapter 3 of Volume I for more information on establishing the level of precision and the target response rate for your survey.)

You do not have to determine either the measurement techniques or any specific analysis techniques that may be needed to meet your research objectives. It usually is best to leave that to the contractor.

The method of analysis used in the CO PEM study was to use the recently-developed miniaturized personal exposure monitors to measure CO in commercial settings in five California cities and suburbs. Then, a number of hypotheses were tested by determining whether there were significant differences between sample results. In all, 588 commercial facilities were visited, including retail stores, office buildings, hotels, restaurants, department stores, and adjacent sidewalk and street intersections. Altogether 5,000 observations were recorded instantaneously at one-minute intervals as the investigators walked along sidewalks and into buildings.

• Step 5: Define the Preliminary Tabulations

At a minimum, you should prepare a list of the preliminary tabulations (table shells) describing the form and content of the tables and graphs you want the contractor to generate when the data file is complete. There is nothing statistically sophisticated about tabulations. They are simply mathematical counts of the number of responses (or specimens) falling into each of several categories that have previously been defined to describe one or more relationships between the variables.

The list of preliminary tabulations should include the title of each table and graph you want the contractor to prepare from the completed data file, and define the horizontal and vertical headings of each. Later, the contractor will total all the responses, specimens, or other items falling under each heading. Note that rarely is it possible to draw up a list of the final tabulations during the planning stage, especially if the subject matter is complex. Usually, most of the tabulations and analyses are not decided on until the results of the data file are in some intelligible and interpretable form.

Let's look at four examples of the tabulations created for the CO PEM study --

(1) Field Survey Dates, Hours, Locations, and Numbers of CO Samples;

(2) Number of Commercial Settings by Type of Setting and Geographic Location;

(3) Statistical Summary of Mean CO Concentrations for Commercial Settings Visited Twice on the Same Date;

(4) Summary of CO Concentrations Collected Simultaneously from Fixed Monitoring Stations and Personal Exposure Monitors.

A slightly abbreviated version of one of the table shells EPA created for the CO PEM study -- the second title listed above -- was the following:

Table 2: Number of Commercial Settings by Type of Setting and Geographic Location

COMMERCIAL SETTING	GEOGRAPHIC LOCATIONS			TOTAL
	Union Square District, San Francisco	University Avenue, Palo Alto	Castro Street, Mountain View	
INDOOR				
Restaurants Hotels Theaters : :				
Subtotals				
OUTDOOR				
Arcade Intersection Midblock				
Subtotals				
GRAND TOTALS				

For additional information, see "Preparing the Preliminary Tabulations" in section A of Chapter 6.

CHAPTER 2

SELECTING THE DATA COLLECTION METHOD

What data collection method should be used for a particular Agency survey? -- There is no general answer and, in many cases, any one of the major traditional collection methods -- face-to-face interviews, telephone interviews, or self-administered mail questionnaires -- may be equally suitable as the primary method.

Researchers no longer arbitrarily consider face-to-face interviews the most effective way of obtaining reliable survey data. If many open-ended questions and extensive probing must be used, it is likely that the presence of a skilled interviewer will motivate the respondents to provide the richest and the most comprehensive data. However, in many other research situations, phone interviews or mail surveys may be just as effective in eliciting the needed data or even more so -- and at a lower cost.

In some cases, the nature and scope of the problems the survey proposes to address may not be defined well enough to begin designing an effective questionnaire and systematically collect data from the target population -- especially when the Agency is dealing with an emerging problem, a new field of science or technology, or a population that has never been studied before. Using exploratory research techniques such as focus groups or in-depth interviews with a few of the potential respondents may be a fruitful way of identifying key topics or hypotheses for subsequent investigation using more traditional statistical measurement techniques.

In the remainder of this chapter we will look at --

- The main characteristics of the methods most often used to collect survey data for EPA;
- The factors that must be taken into account in determining the most appropriate method for a particular Agency-sponsored survey; and
- How to assess the suitability of the proposed collection method(s).

Preceding page blank

A. PRINCIPAL DATA COLLECTION METHODS

This section examines the five most frequently used methods of collecting survey data. First we will look at the three traditional methods used in statistical research and then at two exploratory research techniques that are applicable when the study objectives are not defined precisely enough to begin a systematic data gathering effort.

A combination of collection methods may be used for a major survey. For example, exploratory techniques may be used early on to clarify key topics. Or, if a mail survey is chosen as the primary collection method, telephone or face-to-face interviews may be used later to contact respondents who do not reply within a certain time-limit. A combination of mail and telephone interviewing may be used, whereby respondents are mailed background information and a telephone interview is scheduled later. Telephone interviews may also be used as a back-up for face-to-face interviews after several attempts to contact the respondent in person have failed.

1. Traditional Survey Research Methods

The three most frequently used methods for collecting quantitative (statistical) survey data are --

- Face-to-face interviews
- Telephone interviews
- Self-administered mail questionnaires

QUANTITATIVE
DATA
COLLECTION
METHODS

The data collection instrument for all three traditional collection methods is a "structured" questionnaire. The questions, their sequence, and their wording are fixed in a structured questionnaire. If interviewers are used, they may be allowed some leeway in asking the questions, but generally very little. How much leeway is specified in advance.

- Face-to-Face Interviews

Face-to-face interviewing has been the mainstay of survey research methodology for more than 30 years. It has been used for many EPA surveys during the past ten years. Coupled with a well designed, well tested questionnaire, the face-to-face interview is a powerful, indispensable research tool.

It is adaptable to a wide variety of research situations and is uniquely suited to in-depth explorations of issues.

In a face-to-face interview, selected individuals (the members of the sample) are visited in their homes or workplaces by trained interviewers and asked to respond to a fixed set of questions. The interviewers record the respondents' answers on a printed questionnaire. The answers are the "raw data" that are subsequently processed, studied, and analyzed to solve the problems the survey was designed to address.

- Telephone Interviews

Telephone interviewing is rapidly becoming the principal method of collecting survey data in research situations where probing or in-depth exploration of the issues is not required.

There are two kinds of telephone interviewing techniques: (1) traditional and (2) computer-assisted telephone interviewing (CATI).

= Traditional telephone interviews are similar to face-to-face interviews. The interviewers pose questions to individual respondents at their homes or workplaces by telephone and record the answers directly onto a printed questionnaire. The interviewers generally work from one central location under the supervision of an experienced researcher.

= CATI, on the other hand, is a recent innovation in survey methodology. A printed questionnaire is not used. Instead, researchers program a set of questions onto a computer tape. The interviewer sits in front of a video terminal and reads the questions to the respondents over the telephone as they appear on the screen. The interviewer types the respondent's answers on a keyboard attached to the terminal, and they are automatically entered into the computer.

This radically different interview technique not only speeds up the collection and processing of respondent information, but also avoids the human errors normally associated with handling, checking, and transferring data from a printed questionnaire into machine-readable form.

CATI also has other advantages. It permits the use of very complex "skip" patterns. Depending on the response the interviewer enters, the computer can be programmed to determine which question to present on the screen next. It also provides the interviewer with instant feedback if an impossible or out-of-range answer is entered.

- Self-Administered Mail Questionnaires

  Like face-to-face interviews, self-administered mail questionnaires have been used for several decades to collect survey data. EPA relies heavily on this traditional survey research method to collect complex technical and scientific information from business and industry. In a mail survey, researchers send printed questionnaires to the respondents at their homes or businesses. The respondents complete the forms and return them by mail.

2. Exploratory Research Methods

The Agency sometimes must explore emerging problems dealing with issues about which little is known. We may have determined that a statistical survey is the only way to get the data that will allow us to explore the central issues of these emerging problems, but some aspects of the issues are not defined well enough for us to begin constructing a structured survey questionnaire. In such cases, "unstructured" survey research methods may prove effective in clarifying the key issues.

The two most frequently used unstructured interviewing techniques are --

- Individual in-depth interviews
- Focus group interviews

EXPLORATORY
RESEARCH
TECHNIQUES

Let's briefly examine these techniques.

- Individual In-Depth Interviews

  This exploratory research technique involves individual in-depth discussions with a few individuals in the populations of interest who are knowledgeable about, or involved in, the issues the Agency proposes to study. The interviews will be guided by a topic outline rather than a fixed set of questions characteristic of a structured questionnaire, which is used for virtually all statistical surveys.

In in-depth individual interviewing, probability selection methods generally are not used to choose those who will be interviewed. Instead, a "convenience" sample, representative of different segments of the target population, will be drawn. Any number of individuals may be chosen to participate in the study. The interviewers who are picked to conduct the in-depth interviews must be carefully chosen. They should have experience in conducting in-depth interviews and knowledge of the subject matter.

In-depth individual interviews are particularly valuable when researchers are unsure about (a) which topics are most relevant to the research objectives; (b) whether members of the target population are likely to have the kinds of information the Agency needs; (c) how to phrase certain items on the survey questionnaire; (d) what type of question format is likely to be most effective for obtaining specific information on certain topics (e.g., open or closed questions); (e) which topics the members of the target population are likely to consider threatening or particularly sensitive; and (f) which subgroups in the target population are most likely to be able to supply specific data the Agency needs.

- Focus Group Interviews

Focus group interviews are another valuable "unstructured" research technique featuring informal discussions with individuals selected from the target population. The participants are members of the target population who are called together for discussions focussed on specific issues or specific parts of the proposed survey questionnaire. Focus groups often will unearth aspects of emerging problems that might not surface in individual in-depth discussions. They are especially appropriate for exploring the attitudes, opinions, concerns, and experiences of selected segments of a population of interest; of identifying key concepts; of helping to phrase questions so they will be clear to all potential respondents; and of evaluating drafts of survey questionnaires. Focus groups also may be used early in the development stage of a research project to help the Agency determine whether a quantitative survey is feasible.

As with in-depth individual interviews, probability sampling techniques generally are not used to select the study participants. Instead, several relatively homogeneous groups of six to twelve people are selected at random from various subgroups of the target population. From two to as many as twelve groups may be formed, each led by a skilled moderator knowledgeable about the study objectives. The moderator interacts with the participants and "focuses" the discussion on a few topics of special interest to the researchers.

A topic outline is prepared at the beginning of the study. Usually, fairly general topics are identified for the first group to discuss, then researchers gradually focus the discussions on more specific subject matter. The groups usually meet for about two hours. Although the topic outline is used as a general discussion guide, the participants are given ample opportunity for spontaneous comment -- provided they do not stray too far from the material in the outline.

B. COMPARISON OF THE THREE TRADITIONAL COLLECTION METHODS

Earlier in this chapter we said that no collection method is intrinsically better than any other. However, certain methods are clearly more appropriate in certain research situations and just as clearly contraindicated in others. This section highlights some of the principal distinguishing features of each of the three traditional collection methods.

1. Special Characteristics of Face-to-Face Surveys

Face-to-face interviewing is frequently used at EPA for collecting survey data from the general public. Moreover, it often is the only viable approach for collecting highly complex, sensitive, technical information from business and industry. Because face-to-face interviewing is still the predominant method of collecting data for EPA survey work, much of the discussion in this Handbook pertains to this method.

In-person interviews have many advantages. They generally achieve a higher response rate, greater cooperation, and more complete and consistent data, especially when in-depth exploration of the issues is desirable. Face-to-face interviews are uniquely suited to probing -- a technique used to study the respondent's knowledge of key issues, frames of reference, or, more typically, to clarify and learn the reasons for their answers. The disadvantages of face-to-face interviewing are higher costs and personnel requirements, and the need for extensive training of field staff and close supervision of interviewers throughout the data collection period.

More specifically --

- Face-to-face interviews are the only viable data collection method when first-hand observations of the respondents or the interview site are necessary. Both telephone interviews and mail surveys are inappropriate when eye-witness reports are desirable.

- In some household surveys, particularly when the general public is being interviewed, respondents are more cooperative and give less biased replies if visual aids are used to prompt their answers. Face-to-face interviews are uniquely suited to the use of these aids.

  For example, interviewers can show respondents a calendar to refresh their memories about specific events or time intervals. Or, instead of reading a long list of possible replies, interviewers can hand respondents a checklist (or "prompt card") of suggested answers to elicit an appropriate reply. When an interviewer verbally gives respondents a choice of three or four possible answers, they often have difficulty remembering them all. The net result is a bias towards the first or the last item mentioned. In addition, if interviewers must question respondents about their income or other topics that many people consider too sensitive to discuss with a stranger, prompt cards listing the reply categories tend to cut down on inaccuracies and outright refusals to answer the question.

  Similarly, in a survey of the general public where respondents are required to evaluate a product or other object (a new pollution-control device, for example), face-to-face interviews may be the only viable data collection option. If interviewers are given products for business or industrial respondents to evaluate, however, it may be feasible to mail the firms a sample of the item (or different versions of the product) in advance, and schedule a follow-up telephone or mail interview to get their reports or opinions.

- In face-to-face surveys, smaller, more geographically concentrated samples must be used to hold down costs. Setting up a complex field operation in a large number of sampling areas to interview only a few respondents in each area obviously is prohibitively expensive. To hold down costs, researchers "cluster" respondents in a few selected geographic areas and set up mobile field units to collect the data. Field supervisors remain at a more central location. Clustering does increase the sampling error of the survey, however.

  Widely dispersed samples have little effect on both telephone and mail surveys, on the other hand, because they are generally operated from a centrally-located office.

- Face-to-face surveys are more costly to administer than either telephone or mail surveys. Coordinating, hiring, training, and supervising interviewers and field staff at several locations is complicated. Moreover, the paperwork is much more involved. In addition to the questionnaire, it may be necessary to use as many as 20 different forms and documents to coordinate and control the fieldwork and the processing operations, i.e., confidentiality statements, prompt cards, interviewer calling cards, press releases, interviewer progress reports, interviewer evaluation forms, respondent verification/evaluation forms, and letters giving respondents advance notice of the survey.

2. Special Characteristics of Telephone Surveys

Telephone surveys cost about half as much as face-to-face surveys of comparable size, given the present development of both technologies. They are also easier to manage, produce faster results, and, with few modifications, can be used in most research situations where one-to-one interviewing is indicated.

Some of the advantages of telephone interviews are --

- In Surveys by Telephone, Groves and Kahn report that the overall cost of a telephone survey is 45 to 65 percent lower than a comparable face-to-face survey (see the list of recommended sources at the end of this chapter). Cost savings result from the fact that about one-quarter as many interviewers are needed to reach the same size sample, and the cost of training the interviewers is about one-fifth as much. Moreover, travel costs for interviewers and field staff are virtually nil, and repeated callbacks to the respondents produce no significant cost increases.

- Monitoring, administration, and quality control are simpler than in face-to-face surveys because no far-flung field operation is necessary. Moreover, it is easier to correct interviewer mistakes quickly. People on the contractor's staff who review and edit the completed questionnaires are typically close by, and can provide feedback to the interviewers about errors and omissions relatively quickly. Finally, respondents can easily be recontacted after the initial interview to correct inaccuracies, inconsistencies, and omissions.

- Results can be obtained more quickly from telephone surveys than from either of the other two major collection methods. The interviewing, monitoring, training, editing, and coding operations are usually centralized in one location. If any changes in the questionnaire or interviewing procedures have to be made because of problems encountered in the pretest, the researchers can incorporate them into the main survey quickly. Even after the interviewing in the main survey is under way, it is easy to notify the interviewers immediately about any needed changes. Follow-up interviews to check the interviewers also are much easier.

  If computer-assisted telephone interviewing is used, all the time-consuming manual screening, editing, coding, and data entry operations required for the other data collection methods (including traditional telephone interviewing) are unnecessary.

- With a few modifications, telephone interviews can be used in almost all research situations where face-to-face surveys are suitable.

  For example, if pictures or products must be shown to the respondents to motivate or enable them to answer certain questions, these items can be mailed to the respondents and an interview scheduled at a later date. This combined mail/telephone technique is widely used in marketing surveys.

  The "prompt cards" face-to-face interviewers show respondents to motivate their replies are not possible in phone surveys, of course. However, the questionnaire can be modified to obtain the same information. The most common procedure is to break questions with multiple-choice replies into a series of simpler questions and offer the respondents a set of Yes/No alternatives until all possible answers are covered.

- Telephone interviews permit access to respondents located in areas where face-to-face interviews are especially difficult -- locked apartment or office buildings, subdivisions with security guards preventing access, dangerous neighborhoods, etc.

At the same time, telephone interviewing has several disadvantages.

- Response rates for national telephone surveys remain at least five percent lower than comparable face-to-face surveys despite considerable improvements in interviewer training, feedback procedures, and monitoring techniques during the past few years. The reason is that respondents generally find telephone interviews more tedious and less rewarding than face-to-face interviews, hence they tend to be less cooperative over the phone.

- Telephone interviews are not the best way of collecting factual data if respondents have to search their records or consult with others. However, it is possible to mail respondents background information in advance and schedule a follow-up phone interview later to obtain the needed data.

- Of course, interviewers cannot reach people who have no phones. This means that important subgroups such as low-income people will be underrepresented in surveys of the general public if telephone interviews are used exclusively as the collection method.

  Surprisingly, it is possible to reach people with unlisted numbers in a phone survey. Researchers use a sample-selection technique called "random digit dialing," which features a computer-assisted random selection of telephone numbers. If they simply chose numbers at random from a telephone book, unlisted numbers would be excluded from the sample. About 20 percent of all phone numbers are unlisted.

  Random digit selection has two main disadvantages, however. It is difficult (a) to distinguish between commercial and residential units in the sampling frame and (b) to determine whether units that do not respond are eligible respondents because there is no one on the other end to ask.

3. Special Characteristics of Mail Surveys

Like face-to-face interviews, self-administered mail questionnaires have been used effectively for decades to collect survey data. Mail questionnaires are particularly appropriate for obtaining detailed technical and scientific data and are the least costly of the three collection methods.

More specifically --

- Mail surveys are indispensable for collecting certain kinds of detailed technical data. They are especially appropriate if respondents must consult their records or other people for the necessary data. Self-administered questionnaires allow respondents great flexibility in preparing replies. Respondents have time to think about the questions, gather information from their files, and get advice from others at their own convenience.

- Mail questionnaires are the least costly of the three traditional collection methods, largely because the cost of interviewers is nil or limited to call-backs to assure an acceptable response rate. Moreover, broad geographic coverage is possible with comparatively little effect on the overall cost of the survey.

- Respondents generally are most honest in mail surveys and tend to give fewer "socially-desirable" responses. In an interview survey, particularly in the presence of an interviewer with whom they have established a good rapport, respondents tend to give more socially-acceptable, less critical replies. For example, if respondents are asked if they like living in their community, they tend to say they do, even though on the whole they may dislike it greatly. The same question on a mail questionnaire will elicit more truthful responses.

Mail surveys have some limitations. For example --

- Mail questionnaires must be very carefully designed to compensate for the lack of social interaction that other collection methods provide. Researchers must depend entirely on the questions and written instructions to elicit satisfactory responses and motivate the respondents to cooperate.

The kinds of questions that are suitable for self-administered questionnaires are relatively limited, especially for household surveys. Open questions must be used sparingly. More than a few requests for lengthy answers may result not only in many refusals to answer those questions but also may push respondents to abandon the questionnaire altogether. Generally, if respondents are required to read any but the simplest language or to write out answers in their own words rather than circle or check a printed response, the results tend to be very poor. Of course, these concerns are less likely to be a problem if the respondents are representatives of businesses or industries.

- Mail surveys may be inappropriate if the researchers want respondents to complete the questionnaire with no involvement from others. When questionnaires are self-administered, it is impossible to know the circumstances under which they were completed.

- A substantial follow-up effort is almost always necessary to achieve a reasonable response rate in any voluntary mail survey. To increase the response rate, researchers sometimes give respondents the option of telephoning their replies rather than mailing back the completed questionnaire.

- Compared to other data collection methods, self-administered questionnaires produce a higher number of inaccurate and incomplete responses, largely because no interviewer is present to instruct and motivate the respondents.

C. FACTORS AFFECTING THE CHOICE OF COLLECTION METHODS

A host of interrelated design factors as well as the time and funds available for the survey affect the contractor's choice of the primary data collection method for a particular survey.

In the remainder of this section we will briefly examine the selection factors that normally determine the choice of the primary data collection method for a statistical survey. They are --

- Characteristics of the target population
- Data requirements
- Obligation to reply
- Target response rate
- Time available
- Funds available

MAJOR
SELECTION
FACTORS

1. Characteristics of the Target Population

The characteristics of the target population often are an important consideration in selecting the primary data collection method.

For example, mail surveys of the general public have lower response rates than either of the direct interviewing techniques. Most surveys of business populations, on the other hand, use mail questionnaires as the primary collection method and follow up incomplete or incorrect responses with telephone interviews.

Face-to-face interviews are generally the preferred approach for elderly respondents and those with limited education. Low-income respondents also do best in face-to-face interviews.

The location and distribution of the target population are also factors. Face-to-face interviews are more cost-effective when the target population is concentrated in a small geographic area, such as a particular city or county. If the target population is widely dispersed, however, travel and administrative costs may make a face-to-face survey prohibitively expensive and time-consuming. Self-administered questionnaires or telephone interviews are more realistic options. Mail surveys are least affected by a widely dispersed sample.

2. Data Requirements

The general nature, extent, and complexity of the data requirements are important determinants in choosing the primary collection method. Mail questionnaires should not be used to survey the general public except when answers to a few short questions are needed. Otherwise, it is best to use face-to-face or telephone interviews.

The data requirements of many organizational surveys require respondents to consult their records or other people in order to prepare adequate replies. A self-administered mail questionnaire may be the only feasible way of getting the necessary data in such cases.

If it is necessary to ask a large number of questions respondents may consider "threatening" or unusually sensitive, it is preferable to use face-to-face interviews. To minimize the impact of what may be perceived as potential threats to their operations, business or industrial respondents, for example, may furnish inaccurate or incomplete replies. If it is necessary to collect highly sensitive technical data, the contractor may recommend using trained investigators to make first-hand observations of records or physical facilities to ensure that the Agency obtains complete and valid data.

3. Respondent's Obligation to Reply

The respondent's "obligation" to provide the information the Agency needs often has a critical impact on the choice of the primary collection method. In some cases, the Agency can make responses from businesses and other organizations mandatory. If so, the respondents must provide the required data or face civil or criminal sanctions. Whenever a mandatory response is required, a relatively high response rate is ensured, no matter what collection method is used. Even self-administered mail questionnaires become a viable option. They normally yield so few responses in a voluntary survey they cannot be used for collecting Agency data. (In a survey of the general public, of course, response is always voluntary.)

4. Target Response Rate

The collection method likely to produce the highest response rate given the available funds is obviously preferable. Face-to-face surveys tend to have the highest response rate, other things being equal. However, telephone surveys can produce response rates nearly as high if they are skillfully designed and carried out. As for mail surveys, unless responses are mandatory and considerable follow-up work is done using telephone or in-person interviews, they are unlikely to achieve the 75 percent minimum response rate we recommend for all Agency-sponsored surveys.

5. Available Time

The length of time the Agency can wait to get results also may be a deciding factor in the selection of the data collection method. Computer-assisted telephone interviews have by far the fastest turn-around time. Conventional telephone surveys also can be done more quickly than face-to-face surveys. Mail surveys are generally not appropriate if time is of the essence.

6. Available Funds

The amount of money available for the survey, too, is almost always a critical factor in choosing the primary data collection method.

As we indicated earlier, individual face-to-face interviews are the most expensive way of collecting survey data, other things being equal. Personnel costs (for interviewers, supervisors, trainers, and quality control staff at different field locations) are approximately double that of a comparable telephone survey, where the interviews are usually conducted at one central location. Mail surveys always are the least costly option, largely because the cost of interviewers is limited to some follow-up calls to increase the response rate or to correct inconsistencies and missing or inaccurate replies.

Nevertheless, the least expensive option should not be selected unless it will produce results of acceptable quality. Sometimes it is better to use a higher-cost method and reduce the size of the sample. For example, a mail survey using face-to-face or telephone interviewers to follow up incomplete or unanswered questionnaires usually produces higher quality results than a "pure" mail survey, even if a smaller sample is used to hold down costs.

D. ASSESSING THE SUITABILITY OF THE PROPOSED COLLECTION METHOD

We have recommended that you leave the selection of the collection method(s) up to the contractor. However, as the representative of the sponsoring office, you will have to approve the contractor's choice. The previous discussion of the special features of the three traditional data collection methods and the influence of various survey design factors will help you assess the appropriateness of the proposed method.

To further guide your assessment, Exhibit 1 on the next page indicates the methods most likely to produce satisfactory results under a variety of circumstances.

EXHIBIT 1

GUIDE FOR CHOOSING A DATA COLLECTION METHOD

AGENCY REQUIREMENTS		LIKELY TO BE BEST CHOICE
GENERAL	• Fast turn-around	• Telephone*
	• Lowest possible per unit cost	• Mail
	• Highest possible response rate	• Face-to-Face
	• Fewest possible errors and biases	• Face-to-Face or Telephone
SPECIAL DATA	• Complex technical data (in a mandatory survey)	• Face-to-Face or Mail
	• Detailed data (in a voluntary survey)	• Face-to-Face
	• Respondent's opinion of a product or device	• Face-to-Face**
	• Highly sensitive information	• Face-to-Face or Mail
COVERAGE	• Coverage of all sub-groups in population	• Face-to-Face or Mail
	• Coverage of widely dispersed sample	• Mail
	• Coverage of high-crime or remote areas	• Telephone or Mail
SPECIAL AIDS OR TECHNIQUES	• Extensive probing	• Face-to-Face
	• Third-party observation of records or facilities	• Face-to-Face
	• Respondent diaries	• Face-to-Face or Mail
	• Respondent consultation with others or record searches	• Mail
	• Visual aids (calendars, scales, etc.)	• Face-to-Face or Mail**

* CATI is especially effective.
** Telephone may be satisfactory if visual aids are mailed to the respondents in advance.

Although one of the traditional collection methods will ultimately be selected for testing purposes and for the main survey, using one of the exploratory research techniques discussed in section A may considerably improve the survey design. At a relatively low cost either individual in-depth interviews or focus group interviews often can clarify problems that may be difficult and costly to correct once the survey proper is under way.

FOR ADDITIONAL INFORMATION ON DATA COLLECTION METHODS --

- Mail and Telephone Surveys: The Total Design Method, D. A. Dillman, John Wiley & Sons, New York, NY 1978. Chapter 2.
- Survey Research Practices, G. Hoinville, R. Jowell, and associates, Heinmann Educational Books, London, England, 1978. Chapter 2, "Unstructured Design Techniques."
- Surveys by Telephone, R. M. Groves and R. L. Kahn, Academic Press, Inc., New York, NY, 1976.

CHAPTER 3

DEVELOPING THE QUESTIONNAIRE

A well-designed, thoroughly-tested questionnaire is the most basic tool in survey research. Developing a valid questionnaire for an Agency-sponsored survey requires close collaboration by the sponsors and the contractor throughout the design and testing process.

In this chapter we discuss --

- The principal steps in the development of a good survey questionnaire;
- The respective roles of the project officer and the contractor in designing and testing it;
- How to review drafts submitted for Agency approval; and
- How to monitor the activities designed to pretest the questionnaire.

A. STEPS IN THE DEVELOPMENT OF A SURVEY QUESTIONNAIRE

This section discusses the steps normally involved in developing a structured questionnaire for a statistical survey. The development process we will discuss involves 16 steps, the majority of which are performed by the contractor. Agency-sponsored surveys that are largely repetitions of earlier studies may shortcut many of the steps, but for surveys that address new environmental concerns, a thorough questionnaire-development effort is strongly recommended.

Preparing a survey questionnaire appears to be an easy task, but it is extremely difficult -- even for an experienced questionnaire designer. In no case should you or the contractor begin to draft the questionnaire until the Agency's data requirements have been clearly framed. The reason is that each question must have an obvious link with the data requirements. These requirements then must be transformed into operational concepts and expressed

Preceding page blank

in a logical series of questions, which, when combined and analyzed, will be the measures of those concepts.

Usually several drafts of the questionnaire -- for one or more pretests and for a pilot test replicating the actual conditions of the main survey -- must be prepared and reviewed before a final version is ready to be printed for the main survey. If several versions of the questionnaire have to be designed to accommodate the needs of different types of respondents, more drafts may be necessary.

A summary of the questionnaire-development process we will discuss is given in Exhibit 2 on the next page. The checkmarks indicate the six steps in which the Agency sponsors play the primary role. This role is generally limited to (a) specifying the research topics, (b) reviewing drafts, and (c) monitoring the overall design and testing process. More specifically, as the survey sponsors, you are responsible for --

= Preparing a preliminary analysis plan establishing the research and analytical objectives of the survey (Step 1);

= Supplying the subject matter of the questionnaire, either in the form of a list of topics or a preliminary set of questions (Step 2);

= Overseeing reviews of all draft questionnaires submitted by the contractor and expediting any internal or OMB approvals or clearances that may be required (Steps 5, 7, 11, and 15); and

In addition, you are responsible for monitoring all the questionnaire-development and testing activities the contractor performs, which are covered in Steps 3, 4, 6, 8, 9, 10, 12, 13, 14, and 16.

The discussion of the individual steps, which follows, will cover the Agency's role in specifying the research topics. Section B will show you how to review questionnaire drafts, and section C explains how to monitor field tests of the questionnaire.

We recommend that you use only members of the target population in any preliminary investigations you intend to conduct in preparation for writing the questionnaire, such as --

= Exploratory studies (individual in-depth interviews or focus group studies) to clarify difficult issues or test draft questions you expect to ask (see Step 3);

EXHIBIT 2

THE SPONSORING OFFICE'S TASKS IN THE QUESTIONNAIRE-DEVELOPMENT PROCESS

✓	DETERMINE ANALYSIS REQUIREMENTS	1
✓	DRAFT LIST OF TOPICS OR SUGGESTED QUESTIONS	2
	CONDUCT EXPLORATORY GROUP OR INDIVIDUAL INTERVIEWS	3
	PREPARE FIRST DRAFT OF QUESTIONNAIRE	4
✓	REVIEW/APPROVE DRAFT OF QUESTIONNAIRE	5
	PREPARE PLAN FOR PRETEST	6
✓	INITIATE CLEARANCES FOR PRETEST/MAIN SURVEY	7
	PRETEST ON A SAMPLE OF THE TARGET POPULATION	8
	DEBRIEF INTERVIEWERS/ ASSESS PRETEST FINDINGS	9
	REVISE QUESTIONNAIRE/ PREPARE PLAN FOR PILOT TEST	10
✓	REVIEW REVISED QUESTIONNAIRE AND PILOT TEST PLAN	11
	RECRUIT INTERVIEWERS/ PREPARE TRAINING MATERIALS	12
	PILOT TEST QUESTIONNAIRE AND ASSESS RESULTS	13
	REVISE COLLECTION PROCEDURES FOR MAIN SURVEY	14
✓	REVIEW AND APPROVE PROCEDURES FOR MAIN SURVEY	15
	PRINT QUESTIONNAIRE	16

= Informal pretests to check the content, wording, or format of a proposed questionnaire (see Steps 6-9); and

= A pilot test of the near-final version of the questionnaire and the data collection and processing procedures (see Steps 10, 11, 13, and 14).

Let's look now at the individual steps in the development of the questionnaire.

- ## Step 1: Determine the Analysis Requirements

The first step in constructing a structured questionnaire for an EPA-sponsored survey is to prepare a preliminary analysis plan. Because you, as sponsors of the project, are likely to have greater expertise in the subject matter of the research, you -- not the contractor -- should prepare the analysis plan. As discussed in Chapter 1, the analysis plan should define (a) the purpose of the survey, (b) the research objectives, (c) the key variables, (d) the analytic approaches and methods to be used to achieve the stated objectives, and (e) a list of preliminary tabulations, which will allow you and the contractor to decide which types of analyses will best reveal the full informational content of the data base. You should include at least a draft analysis plan in the Request for Proposals (RFP) the Agency issues for contract support for the survey. Then, after a contract is awarded, the contractor can refine the draft and submit it for approval along with the other components of the work plan.

- ## Step 2: Draft a List of Topics or Suggested Questions

We suggest that you prepare a comprehensive list of research questions and, perhaps, an informal list of the items you would like to see on the final questionnaire. Keep in mind that all questionnaire items must be clearly relevant to the informational and analytical objectives of the research. Questions should not ask for information that may be "nice to have." If you decide to draft an informal list of questions, therefore, as you write each item, ask yourself, "Why do I want to know this?" "It would be interesting to know" is not an acceptable response. Also, don't attempt to write the questions verbatim or to format the questionnaire. It's best to leave those tasks to the contractor (see Step 4).

Before preparing your list of research topics and preliminary questions, we suggest you look for questions or scales that have been used in earlier Agency surveys to explore various environmental issues. In addition, you may find questions or scales used in other survey reports helpful in framing your research questions.

A search of this type may seem time-consuming and tedious, but it often is time well spent. Even if you find only a few good items, this may cut down on the time required to test the questionnaire. Moreover, the search generally will give you a better perspective on your analysis needs.

If you do find some usable questions, they are unlikely to cover all aspects of the problems the new survey is intended to address, especially since EPA often tackles evolving issues on which little research previously has been done. No doubt, you will have many new questions you expect the contractor to explore.

Keep in mind that any list of topics or questions prepared at this stage should be regarded as preliminary. Only after exploratory studies and one or more advance tests of the data collection instrument are completed can you be reasonably confident of having a questionnaire that will meet your data and analysis objectives. Some compromises in the data requirements may be necessary if it turns out that respondents are unable or unwilling to answer certain kinds of questions.

- Step 3: Conduct Exploratory Interviews with a Few Individuals in the Population

Even if you succeed in preparing a reasonably complete list of topics or preliminary questions, you may find that there are still gaps in your understanding of the issues. If so, before the contractor begins to draft the initial draft of the survey questionnaire, it may be productive to explore some of the key issues with a few members of the populations you plan to investigate.

A series of focus group interviews or in-depth interviews may prove fruitful in resolving uncertainties at this early stage of the questionnaire's development. To date, EPA has not used either of these exploratory research techniques extensively, but other agencies have found them highly effective in resolving a range of conceptual problems that would be prohibitively costly or impossible to resolve later in the development of the questionnaire. Individual in-depth interviews or focus groups can be used to explore attitudes, opinions, concerns, and experiences of potential respondents; develop data specifications; test the wording of questions; or even to evaluate an entire draft of a questionnaire.

These techniques are suitable for exploring issues relating to household or non-household surveys. For example, sometimes it is essential for the sponsors to know the record-keeping practices of the industries they intend to survey so they can determine what kinds of questions the respondents may reasonably be expected to answer. Either of these exploratory research techniques is likely to add two to six weeks to the overall development process. If an OMB clearance is necessary, it may take somewhat longer. However, because the final questionnaire undoubtedly will require fewer refinements and less testing, you may well be able to recover this time before the main survey begins.

Be sure to check with your office's Information Management Coordinator regarding the need for an OMB clearance for this preliminary interviewing work. Sometimes a clearance is necessary.

● Step 4: Prepare First Draft of the Questionnaire

Building on (a) the data and analytic requirements you formulated in Steps 1 and 2; (b) the findings of the exploratory interviews, if any (Step 3); and (c) other specifications in the work plan concerning the data collection, processing, and analysis procedures, the contractor can begin to draft the questionnaire. A structured questionnaire typically consists of --

= Introductory information explaining the objectives of the survey and the reasons the respondent's cooperation is solicited. (In a self-administered questionnaire, this information is usually stated in a cover letter.)

= Identification and control information showing the name of the survey sponsor, the name of the organization collecting the data, the authority for collecting the data (e.g., any applicable statutes), the OMB control number and expiration date of the clearance, various code numbers identifying the individual response unit (the household, business, individual, etc.) and where the unit is located, and any additional information needed for control purposes.

= A set of standardized questions addressing the research problem;

= Instructions to the person filling in the data;

= Definitions of all technical and unusual terms.(An EPA-sponsored survey of businesses or industries frequently will include an entire section on definitions.)

In most cases, once you have formulated the basic content of the questionnaire and approved the work plan, it is best to let the contractor construct the questionnaire. The content and wording of the individual items as well as the overall organization and format of the questionnaire will be major factors in determining whether the survey ultimately produces timely, reliable, useful information.

The questions must be worded so they can be clearly understood, arranged in the best possible order, and capable of eliciting objective, unbiased answers. If the questionnaire is to be self-administered, it has to be designed in a way that will motivate the respondents to make the necessary effort to retrieve, organize, or report the required information in the specified format. If it is to be administered by a trained interviewer, the design and format should facilitate the work of the interviewers in asking questions and recording responses. The format should also expedite the coding and data entry operations during the processing phase.

- ## Step 5: Review and Approve First Draft of the Questionnaire

Extensive reviews of the first draft of the questionnaire (and all subsequent drafts) submitted for Agency approval are vital to ensure that --

= The content is relevant to and focused on the research objectives;

= The wording is clear and unambiguous: and

= The overall organization and format of the questionnaire will facilitate the data collection, processing, and analysis activities.

As project officer, one of your principal responsibilities during the development process is to ensure that the questionnaire is constructed so that it will achieve the objectives of the study. Criteria for a systematic review of draft questionnaires are given in section C; therefore, we will not elaborate further here.

In addition to circulating drafts to key people on the project staff, you should have computer programmers, systems analysts, and statisticians review them as well as people outside EPA who are knowledgeable about the subject matter or the intended uses of the data. After the contractor incorporates changes in the draft, make sure the comments of all reviewers are accounted for.

- Step 6: Prepare Plan for Pretest

While the Agency is reviewing the initial draft of the questionnaire, the contractor should prepare a plan to pretest it informally on one or more subgroups of the target population.

The pretest plan should cover (a) the scope of the test (whether the entire questionnaire or only certain questions will be evaluated); (b) the size and composition of the test sample; (c) the techniques to be used in administering the test (e.g., face-to-face or telephone interviews); (d) procedures for training the interviewers and observers; (e) procedures for conducting and evaluating the test; and (f) the kinds of tabulations and analyses that will be done.

Pretesting is essential for all structured questionnaires, regardless of the data collection method proposed for the survey proper. The techniques used to pretest an interview survey and a mail survey are quite different, however.

= For a face-to-face or telephone survey, one or more informal pretests are mandatory. Rigorous analytic techniques normally are not used, however. Instead, interviewers, observers, and respondents subjectively evaluate various aspects of the questionnaire. At a relatively low cost, pretests can determine whether changes in the wording of the questions, their sequence, or the length of the questionnaire are likely to improve the quality of the survey data. Pretests also may indicate a need for adding or eliminating certain questions.

Usually the contractor will do a few informal tests; then, when the wording and format of the questionnaire have been refined, they will conduct a formal test, called a "pilot test," to evaluate the data collection procedures as well as the questionnaire. For a major interview survey, a full-scale pilot test should be done. (Step 12.)

Some of the techniques used to evaluate pretests of an interview survey are (a) observations by trained supervisory staff; (b) discussions with respondents immediately after the questionnaire is administered; (c) daily interviewer debriefings; (d) interviewer records of call-back rates and the duration of the interviews; (e) tape recordings of a few test interviews; (f) written reports by interviewers on the difficulties encountered in collecting the data, and suggestions for improving the questionnaire, control forms, or the interviewing procedures; (g) debriefings at the conclusion of the pretest with the interviewers, questionnaire designers, field supervisors, and observers; and (h) preliminary tabulations of the pretest data.

= Techniques for pretesting a mail survey tend to be more formal. Usually, a draft of the questionnaire is mailed to a small subset of the target population. The results are then tallied and evaluated.

A less formal method of testing a mail questionnaire is to mass-administer it to a group of respondents "classroom-style," with a moderator and several observers in attendance. Some face-to-face interviews also may be used for testing mail questionnaires at an early stage of their development.

When the contractor submits the pretest plan for Agency review, make sure (a) the pretest sample adequately represents all important subgroups of the target population, (b) the size of the sample is adequate for a valid test, (c) the test conditions approximate those of the actual survey, and (d) enough time has been allowed to analyze the test results and incorporate any necessary revisions in the questionnaire.

Submit the plan along with the questionnaire to approval authorities in your office. If you need to apply for an OMB clearance for the pretest, also have the Information Management Branch of the Office of Standards and Regulations review the plan and the questionnaire at this time.

- **Step 7: Initiate Clearance Request for the Pretest**

Obtaining OMB clearance(s) for all pretests and the main survey in a timely way is a major responsibility of the project officer if data are to be collected from ten or more members of the public. Clearance is mandatory per the Paperwork Reduction Act. The purpose of the OMB review is to ensure that (a) the information that agencies propose to collect is in the public interest, (b) the reporting "burden" (the length of time it takes a respondent to complete a questionnaire or be interviewed) is reasonable, and (c) certain statistical standards are met.

You may submit a clearance request for the pretest (or a series of pretests) along with one for the main survey. Allow two weeks for each Agency office that must review the clearance package before it goes to OMB. Allow a minimum of two months to obtain OMB clearance after you secure all necessary internal approvals. (See Chapter 7 of Volume I for more information on OMB clearance procedures.)

- **Step 8: Pretest on a Sample of the Target Population**

While awaiting the OMB clearance, the contractor sometimes will organize and train the interviewers and other staff to be used for the pretest, but usually it is best to wait until the clearance is granted.

The contractor's principal responsibilities in preparing for the pretest are --

= Selecting the agreed number of respondents from the target population. For an informal pretest, 20 to 50 respondents usually will suffice. Generally, a "purposive" sample rather than a probability sample is drawn so that all subgroups in the target population or specific subgroups of concern are represented.

= Choosing interviewers for the test. Some survey research firms maintain an experienced team of interviewers solely for pretests. Others use only supervisors so they can gain experience that will be useful in training and overseeing the interviewers picked for the main survey. Still others use interviewers with education and experience similar to that of the interviewers to be used for the main survey. In all cases, it is best to use as many interviewers as possible, provided each of them has a sufficient workload to justify the cost of their training and travel.

- Selecting and training one or more field supervisors to oversee the interviewing.

- Training the interviewers in the general purposes of the survey and the specific objectives of the pretest. This kind of training is vital for all the interviewers who participate in the test -- even the most experienced. If the interviewers do not have a thorough understanding of the questions, it will be impossible for the questionnaire designers to determine whether problems with the questionnaire are due to poor interviewing or to the data collection instrument itself.

  The interviewers also must be thoroughly trained in the proper way to administer the questionnaire (e.g., not to arbitrarily reword questions; how to probe and ask other questions when respondents' first answers are inappropriate, inaccurate, or incomplete).

The pretest itself frequently is conducted under conditions similar to that planned for the main survey.

As for the project staff's responsibilities once the pretest is in progress, we recommend that --

- You or members of your staff observe several pretest interviews to gain first-hand experience in how the questionnaire works in practice. Discussions with respondents following each pretest interview -- a major feature of informal pretests -- provide important feedback to questionnaire designers on how respondents interpreted various questions; difficulties they experienced in replying to certain items; how they would ask certain questions; or their feelings about questions to which they responded "Don't know," etc.

- You attend some of the daily debriefings with the interviewers. The purpose of these debriefings is to get immediate feedback from field personnel on problems they have had with the questionnaire so the contractor can make on-the-spot refinements for testing during the next day's interviewing. The interviewers may discuss (a) difficulties they encountered in locating respondents; (b) questions that embarrassed respondents or otherwise made them feel uncomfortable; (c) which items respondents refused to answer and the reasons given for the refusals; (d) difficulties they had in maintaining rapport with respondents; (e) whether the respondents became impatient or bored; (f) whether respondents seemed to want to rush through any part of the questionnaire, particularly the ending; (g) whether the format of the questionnaire was particularly hard to follow; (h) whether any items required further explanation;(i) how long the interviews took; and (j) if there was enough space to record answers, especially to open questions.

(See section C for suggestions on monitoring pretests.)

- Step 9: Debrief Interviewers and Assess Pretest Findings

When the pretest is over, the contractor generally will hold one or more debriefing sessions with all the interviewers, supervisors, and observers who have participated in the pretest.

You and other members of your staff should attend these sessions so that any necessary changes in the questionnaire or training procedures can be jointly agreed to and quickly implemented. The format of these sessions generally is similar to that of focus group discussions (see section A of Chapter 2).

Based on the outcome of the final debriefings and any preliminary tabulations, the contractor will be in a position to determine if further revisions or tests of the questionnaire are needed.

The contractor should revise the questionnaire after each pretest until all problems are resolved. In a major survey, another pretest should be done after each revision because the revisions may cause new problems.

Note: Steps 10-13 may be omitted if no further tests are planned.

- Step 10: Revise Questionnaire and Prepare Plan for the Pilot Test

If you propose to survey more than 500 people (or units), the last step in the testing process for an

Agency-sponsored survey should be a full-scale pilot test -- a more formal type of pretest. A pilot test is, in effect, a "dress rehearsal" for the main survey. Normally, it should duplicate the field procedures as closely as possible, and the questionnaire should appproximate the one that will be used in the main survey.

The first step in preparing for the pilot test is to develop a planning document clearly delineating the objectives of the test. Pilot tests can be used to --

= Evaluate the wording, content, and format of the questionnaire, and test alternative versions, if necessary;

= Identify and correct weaknesses in the proposed interviewing procedures -- the interviewer's instructions and training manuals, the length of the interviews, and the logistics of the field operations;

= Provide a realistic body of data to test the proposed processing procedures -- the specifications and instructions for coding, data entry, computer editing, and tabulation operations.

If the test is carried through to the analysis phase, the preliminary tabulations can provide a final check on the analysis plan.

The time required to conduct, process, and evaluate the results of a pilot test is considerably longer than for an informal pretest. From five to ten months may be required for the pilot -- after the Agency approves the questionnaire. This includes the time required to obtain OMB approval (up to 90 days).

In a pilot test of a face-to-face survey, usually at least 50 respondents and several interviewers at different skill levels are used. It is not unusual to have up to 300 respondents and as many as 20 interviewers. Potentially "difficult" respondents or "hard-to-reach" population groups should be included.

The interviewers also must be selected and trained in the specifics of the test and one or more field supervisors appointed to keep track of the interviewers' workload and evaluate their performance.

- **Step 11: Review Revised Questionnaire and Pilot Test Plan**

You and your staff should critically review the pilot test plan, giving special attention to the proposed tabulations and analyses. Circulate it to computer programmers and system analysts, if necessary.

The contractor should allow enough time to analyze the data and apply the findings before the main survey begins. Important benefits of pilot tests frequently are not realized because the analysis is not planned in enough detail, and insufficient time and resources are committed to it.

If you have not yet applied for OMB clearance of the pilot test, you must do so at this time. We recommend that you combine it with the clearance request for the main survey so the contractor can proceed with the main survey as soon as the pilot test results are analyzed.

- **Step 12: Recruit Interviewers and Prepare Training Materials**

The quality of the interviewing in the pilot test and the survey proper will be greatly influenced by the amount of care taken in selecting and training the interviewers.

As we have seen, a great deal of effort typically goes into the development of the questionnaire so it will effectively yield valid, unbiased data. To achieve satisfactory results in an interview survey, the data must be collected in a systematic, uniform manner from all the respondents.

The interviewers selected for the pilot test usually work in the main survey as well. If the contractor has a permanent field staff in the sampling areas, there probably will be no need to recruit new interviewers. Most large survey research firms maintain a permanent cadre of interviewers located throughout the United States. Having a permanent interviewing staff does not guarantee the quality of the fieldwork will be high, but experienced interviewers are far more likely to collect good data than a group of new interviewers recruited solely for one survey.

In addition to selecting the interviewers, the contractor must (a) develop procedures and materials for training the interviewers and a field supervisor, (b) determine how many training sessions will be needed, and (c) where the session will be held. This can be done while awaiting the OMB clearance for the pilot.

Interviewer training for the pilot test should cover the objectives of the survey, the content and concepts of the questions, interviewing techniques, the procedures to be used to control the quality of the field work, and practice interviews. Instruction manuals and other training materials also should be prepared so their effectiveness can be assessed before the interviewers for the main survey are trained. (See section A of Chapter 5 for detailed information on training.)

- **Step 13: Pilot Test Questionnaire and Assess Results**

Once the interviewers are recruited and trained, the interviewing phase of the pilot test should proceed much like any other data collection operation using a structured questionnaire. The techniques used to observe and evaluate the test are similar to those used in informal pretests (see Steps 8 and 9) with one major difference -- a greater focus on statistical evaluation of the data.

For example, debriefing sessions with all the interviewers and observers are held following the test. The debriefings may alert the analysts to problems with specific questions, the order of the questions, or the length of the questionnaire. As a result, it may be necessary to change or discard certain questions. If the average length of the interviews is too great, some questions may be dropped to stay within the established time and budget constraints -- even if nothing is wrong with the questions.

To carry a pilot test to its logical conclusion, the analysis of the pilot test data should be sufficient to allow the contractor to assess the validity of the analysis plan.

- **Step 14: Revise Questionnaire and Collection Procedures for Main Survey**

When the pilot test is concluded, the questionnaire ordinarily should require few revisions. By gradually fine-tuning the data collection instrument -- through discussions with respondents, interviewer debriefings, observation and monitoring of the interviews, interviewer reports and assessments, data validation, and the analysis of the pilot test data -- the contractor should be in a position to begin the main survey with clear assurance that the resulting data will meet the Agency's objectives.

In addition to modifying the questionnaire, the contractor should submit a revised data collection plan to the Agency for approval before the survey proper begins. The plan should include (a) provisions for training and supervising the interviewers, (b) "rules" for respondent eligibility (respondent rules), (c) rules for following up the initial contacts with respondents, (d) rules for verifying and evaluating the interviews, and (e) the quality-control measures that will be used to ensure that the target response rate for the survey and the response rates established for individual items are achieved. (See section A of Chapter 5 for detailed information on preparing for the interviews.)

• Step 15: Review and Approve Procedures for the Main Survey

The final draft of the questionnaire and the proposed data collection procedures should be critically reviewed by the project staff, data processing specialists, and systems analysts. We strongly recommend that you have a survey expert review these materials (whatever collection method is planned) before granting approval to proceed with the survey. A successful field operation requires close coordination and monitoring by the contractor and the EPA project staff, effective interaction between the interviewers and the respondents, and careful training and supervision of dozens of interviews at several field locations.

If you have not submitted the OMB clearance request for the main survey, do so at this time after clearing it with your Office Director and OSR's Information Management Branch.

• Step 16: Print Questionnaire

The questionnaire for the main survey should not be printed until the results of the pilot test indicate there are no more serious problems. In no case should it go to the printer until you have received an OMB control number. Both the number and the expiration date of the clearance must appear on the form.

Make sure that the contractor orders enough questionnaires. It is best to get 50-100 percent more than the number of respondents. The extra copies can be used for training purposes and practice interviews. Copies sometimes are lost during the distribution process and others are wasted in the field.

Check proofs of the questionnaire received from the printer for spelling and typographical errors. When the printed version arrives, batches should be spot checked for poor print quality, missing pages, etc.

B. REVIEWING DRAFT QUESTIONNAIRES

This section provides instructions for systematically reviewing a survey questionnaire. The instructions are intended to help you critique drafts submitted by the contractor for Agency approval during the development process, as shown in Exhibit 2.

The instructions are presented in three parts. We recommend that you first review (1) the form, content, and wording of each question individually; then (2) the content and organization of the questionnaire as a whole; and, lastly, (3) the overall format.

A checklist of the suggested criteria for this three-stage review is given in Exhibit 3. Use it, along with a copy of the analysis plan (see Chapter 1), to guide your reviews. Also, be sure to circulate review drafts to others with expertise in questionnaire design, data processing, and statistical analysis, as appropriate.

1. Reviewing Individual Questions

Begin your review of the questionnaire by critically examining each question. Review the --

- Form
- Content
- Wording

REVIEWING
INDIVIDUAL
QUESTIONS

- Form

You'll want to look first at the appropriateness of the form -- the answer format -- of each question.

EXHIBIT 3

CRITERIA FOR REVIEWING SURVEY QUESTIONNAIRES

INDIVIDUAL QUESTIONS

- Form
- Content
  - = Relevance
  - = Reasonableness
  - = Sensitivity
  - = Completeness
- Wording
  - = Clarity
  - = Simplicity
  - = Absence of (unintentional) leading or "loaded" terms

GENERAL CONTENT AND ORGANIZATION

- Scope of the questions
- Order of the questions
- Explanatory and control information
  - = Introductory explanations
  - = Instructions
  - = Definitions
  - = Interviewing aids
  - = ID and control information
  - = Data processing directives

OVERALL FORMAT

- General appearance
- Length
- Placement
  - = Of the questions
  - = Of the instructions

There are three reasons: (a) Survey questions are classified by their answer format, (b) the form is the most immediately visible aspect of a question, and (c) the proposed form of the question may affect your review of the content and wording.

To assist you, we'll briefly outline (1) the basic types of survey questions and (2) the advantages and limitations of each.

= Types of survey questions.

There are three basic types of survey questions:

(1) Closed (or closed-ended) questions offer respondents a choice of two or more response options, the most common of which are "Yes/No" and "Agree/Disagree." Sometimes a third option, "Don't know" or "Undecided," is used. Closed questions are sometimes called "fixed alternative," "fixed choice," or "poll" questions. Also classified as closed questions are so-called "multiple-choice" questions, which permit respondents to choose their answer(s) from several response categories.

(2) Open (or open-ended) questions give respondents a frame of reference but permit them to reply in their own words. Traditional open questions allow respondents to give their opinions fully, in language comfortable to them, without restriction. However, open questions do not necessarily call for a lengthy response. They are often used when very short numerical answers are sought -- age in years, expenditures in dollars, volume in cubic feet, etc.

Open questions are further classified as fully-open (the traditional open question) or partially-open. When a question is fully-open, the interviewer simply records the reply verbatim. The questionnaire will include a blank space for the interviewer to write in the respondent's answer. If the interviewers find it necessary to use probes to encourage a more complete answer, they are expected to indicate directly on the questionnaire where they intervened to seek clarification -- usually by placing an "X" after the respondent's reply.

Partially-open questions, on the other hand, are more like closed questions. They appear to be open to the respondent, but they actually provide a fixed set of response options. The interviewer selects the response option(s) closest to the respondent's answers or, sometimes, will guide the respondent to an answer within certain limits. Partially-open questions on self-administered questionnaires provide several fixed response options as well as an "Other-Specify" category.

(3) Scale (or ranking) questions permit respondents to rank their responses according to (a) preference or interest, (b) degree of agreement or disagreement, or (c) some other scale of measurement. Scale questions are actually a special form of closed questions.

= Advantages and Limitations.

Many survey research firms have a decided preference for closed questions. There are three reasons: (1) closed questions tend to be more reliable; (2) they are easier for interviewers, coders, and analysts to deal with; and (3) unlike open questions, they generate no irrelevant, unintelligible responses to complicate the data processing and analysis phases.

Nevertheless, closed questions have certain disadvantages. The major problem is their superficiality. A questionnaire containing only closed questions doesn't get to the heart of issues.

Closed questions also tend to force replies. Sometimes respondents choose any answer to conceal their ignorance about the topic or they may pick a response that does not reflect their true opinion just because they feel compelled to check or circle one of the fixed responses.

Carefully constructed and used in combination with open questions, however, closed questions can be very effective.

Open questions have many advantages. They put a minimum of restraint on respondents' replies and the manner in which they express them. The open format permits interviewers to probe the respondents' knowledge of a subject and their frames of reference, and to clarify or ascertain the reasons for the answers they give. To learn the respondents' true intentions, beliefs, feelings, or attitudes, some open questions should be used. The open format is an invaluable tool for exploring a topic in depth. It is an absolutely essential tool if you are beginning work on a new research topic and need to explore all aspects of the subject.

But open questions are also appropriate when the potential responses are both nominal in nature and sizeable in number, e.g., questions asking for a single-word response such as the respondent's age or income.

The richness of data the open format yields, however, can be a disadvantage when it comes time to summarize the data in concise form. Reducing a large number of varied responses to a few categories that can be treated statistically is a major challenge for coders during the processing. Coding a complex set of open responses is not only time-consuming and costly, but also introduces some amount of (coding) error. If the data categories are extensive, the contractor must develop complex coding instructions, train staff in the proper use of the codes, and make periodic reliability checks to estimate the amount of coding error. (See Chapter 6 for more information on coding.)

There are other disadvantages. Open questions take more time to answer than closed questions. This tends to increase the response burden of the survey. They also require greater interviewer skill in recognizing response ambiguities, and in probing or drawing out respondents, particularly those who are reticent or not highly verbal, to make sure their answers are codable. This aspect of the open format has made some researchers wary about using it except in situations where they are sure of getting well-trained, well-supervised interviewers.

Scale questions are good for measuring attitudes and values because they allow researchers to identify the intensity of respondents' feelings, beliefs, or preferences. You might devise an intensity scale, for instance, to measure a community's preference of air quality strategies.

To help you assess question forms, we conclude our discusssion with a few tips from Sudman and Bradburn's Asking Questions (see list of references at the end of this chapter.)

(1) Open questions should be used sparingly -- for developmental work, to explore a topic in depth, and to obtain quotable material. Closed questions are more difficult to construct but easier to analyze and less subject to interviewer and coder variance.

(2) When lists are used, complete information can be obtained only if each item is responded to with a "Yes/No," "Applies/Does not apply," "True for me/Not true for me," and the like, rather than with instructions such as "Circle as many as apply."

(3) Rating-scales with more than four or five verbal points should not be used. Numerical scales are preferable if more detailed measurement is desired.

(4) Respondents should not be asked to rank their preferences among a number of options unless they can see or remember all the options. In face-to-face interviews where prompt cards are used, respondents can rank no more than four or five options. If many options are present, respondents can rank the three most desirable and the three least desirable. In a telephone interview, rankings can be obtained by a series of paired-comparison questions. However, respondent fatigue limits the total number of alternatives that can be ranked.

- Content

Next, you'll want to review the content of the individual items. Each question should be (a) relevant to the Agency's informational or analytical objectives, (b) reasonable, given the respondents' probable knowledge and experience, (c) sensitive to the respondent's self-interest, and (d) complete. More specifically --

= Relevance.

Each question should be clearly relevant to the informational and analytical objectives of the survey, as defined in the analysis plan. Except for the first one or two questions, which may be designed simply to orient the respondents or put them at ease, each item on the questionnaire should yield a particular piece of data that will contribute to the informational objectives of the survey. Of course, more than one question may be needed to get a complete perspective on a single research question or variable.

= Reasonableness.

The question should ask for information the respondents can reasonably be expected to provide, given their probable knowledge and experience. The extent to which people can respond to the question will affect both the quality and quantity of their responses. Rather than admit their ignorance, respondents may give a false reply or no reply at all.

In reviewing the question, therefore, consider the difficulty of the question from the respondent's perspective.

For example,is the respondent required to recall events or transactions that happened weeks or months ago? According to the Sudman and Bradburn reference mentioned earlier, periods of a year (or sometimes even more) can be used for highly salient topics such as the purchase of a new house, the birth of a child, or a serious auto accident. Periods of a month or less should be used for items with low saliency such as the purchase of clothing or minor appliances.

If detailed information on frequent behavior of low salience is required, respondents can be asked to keep diaries. Diaries will provide more accurate results than memory. In a business survey, the use of records (if available) and direct observation by interviewers will improve reporting of the desired information. In addition to diaries, records, and direct observation, other techniques can be used to motivate respondents to supply accurate data, e.g., (a) probes or follow up questions, (b) verbal reinforcement by interviewers, and (c) interviewing aids such as pictures, calendars, checklists or prompt cards.

- Sensitivity.

In addition to being unable to answer, the respondents may not want to reply to a particular question because they feel some harm may come to them, or they will be embarrassed, or that the information is too personal to divulge to others. The net result is the same as for unreasonable items -- many inaccurate or missing responses.

Therefore, in reviewing the content of individual questions, it is important to consider the sensitivity of each question. Topics many people regard as sensitive are income, assets, profit, religion, political affiliation, and beliefs. Any question dealing with such topics must be well justified. (OMB, in fact, requires additional justification for questions that are likely to be considered intrusive or damaging to respondent self-esteem.)

If the question is not essential, it may be best to drop it. If it is essential, there are ways of minimizing the possibility of inaccurate or missing responses:

(1) Careful placement helps. Locating a question on a sensitive subject towards the end of the questionnaire or grouping it with related questions of a non-threatening nature tends to improve the reliability of the response. (See Placement at the end of this section.)

(2) For obtaining information on frequencies of socially-undesirable behavior, open questions are better than closed questions, and long questions are better than short questions.

(3) If respondents are being asked to rank attitudes or behavior, the scale should start with the least socially-desirable response options. Otherwise, the respondent may choose a socially-desirable answer without hearing or reading the entire set of responses.

(4) In asking about socially-undesirable behavior, it is better to ask respondents whether they have ever engaged in the behavior

before asking them about their current behavior. Also, it is better to ask about "current" rather than "usual" behavior.

= Completeness.

Obviously each question should have all the elements necessary to get the desired information. There are several tests you can apply to each question to determine whether it is complete. For example --

(1) If the respondent is to check only one of a set of fixed response categories, the categories must be exhaustive, i.e., they must cover all possible alternatives. If not, then an "Other-specify" category should be added. Response categories also must be mutually exclusive, i.e, there should be no overlap to confuse the respondent.

(2) If the question contains a time reference, the period or date should be specified.

(3) By the same token, if you want the respondent to reply with a numerical amount, clearly indicate the desired units, such as days, tons, or dollars.

(4) If the respondent is asked to give an opinion on a particular issue, a "Don't know" or "No opinion" response category may be needed. Including such a category frequently will have an effect on the results. Whether or not to include an additional response option of this type depends on how desirable you believe it is to get the respondent's opinion -- even though he or she may have little knowledge of the issues.

(5) Questions should be phrased so that the analysts can distinguish between no response and a response of "Zero" or "None." For example, if an item such as --

Annual volume of chemical waste products

__________ (metric tons)

is left blank, it will not be clear to the analysts whether the firm's waste products total zero tons or whether they simply did not answer the question. This can be remedied by changing the item to --

Annual volume of chemical waste products

☐ None or ____________.
(metric tons)

- Wording

The last set of review criteria for individual questions concerns wording. Each question should be (a) clear and unambiguous, (b) simple and specific, and (c) free of any unintended leading or "loaded" language.

In reviewing the wording, read each question slowly, preferably aloud, and assess its --

= Clarity.

To keep response errors and biases to a minimum, each question should be clearly and unambiguously worded so there is no way for anyone in the sample to misinterpret it.

Words that can change the entire meaning of a question if they are not correctly interpreted should be italicized or underscored. For example, any change in the frame of reference from a previous question should be clearly indicated -- a request for "total gross sales last month," rather than a request earlier in the questionnaire for "total gross sales last year"; or or "monthly net income," rather than "monthly gross income." If necessary,the question should be reworded to eliminate any chance of misinterpretation.

Words with multiple meanings are especially problematic. For example, in a question like "Do you think EPA has treated the chemical industry fairly?" -- "fairly" could mean "justly," "equitably," "not too well," "impartially," or "objectively."

Any unusual words should be defined. (See Definitions later in this section.) Slang and colloquialisms should be avoided -- not because they violate good usage, but because many respondents may not know what they mean.

= Simplicity.

Simply worded questions also help to reduce the number of inaccurate and missing responses. Compound questions giving two or more frames of reference -- so called "double barreled" items -- confuse respondents and result in many invalid responses. A question like "Do you feel that air pollution is a serious problem and that dust from construction sites is the major cause?" would confound many respondents, who may agree with only half the question. The classic example of a double-barreled question is "Have you stopped beating your wife?"

Making questions as specific as possible tends to make the respondent's task easier, which, in turn, results in fewer invalid replies. Normally, a question should tap a specific opinion, not a general attitude. Items should be directed to specific rather than general concerns.

= Absence of leading or "loaded" terms.

Respondents generally want to be thought of as good people. Even in circumstances where they might be expected to be strongly opposed to something or someone, respondents tend to choose an answer that is most favorable to their self-esteem, that they think makes them look intelligent or thoughtful, that they think the interviewer would like them to give, or that is in accord with social norms. A further factor leading to bias is a desire to be polite to an interviewer, who usually is a stranger. In being polite, respondents will hesitate to say unkind things they believe might offend the interviewer.

Therefore, any question asking about socially-desirable or socially-undesirable behavior or attitudes tends to produce bias and must be worded with care. One of the most common traps questionnaire designers fall into, in fact, is to use leading or "loaded" words, particularly words that are loaded with "social desirability."

At the same time, there are instances where it may be desirable to use leading questions.

For example, you might ask the question, "When was the last time your exhaust filtration equipment failed to function properly?" The equipment actually may never have failed. On the other hand, if the researchers believe the respondents may have a tendency to underreport such failures, asking the question this way may result in more accurate statistics.

2. Reviewing the Overall Content and Organization

Next, examine the questionnaire as a whole, specifically looking at the --

- Scope of the questions
- Order of the questions
- Explanatory and control information

OVERALL
CONTENT AND
ORGANIZATION

Scope of the Questions

The questionnaire should, of course, cover all aspects of the problem. Since you, as the survey sponsors, undoubtedly will have contributed the basic substance of the questionnaire, your review of the overall content at this point should be a simple matter of making sure that the draft encompasses all the Agency's data requirements. The analysis plan will be invaluable for guiding this part of your review.

Order of the Questions

Questions should be logically ordered and grouped into coherent categories. The categories do not necessarily have to be labeled, but similar items should be grouped together. A transition statement should mark significant change in topics.

Whether respondents complete the questionnaire on their own or in the presence of an interviewer, they are less likely to become fatigued and will make fewer mistakes if they don't have to shift mental gears constantly. Most respondents are not experts at questionnaire design, but they certainly can distinguish between a questionnaire that is well organized and one that is poorly ordered, duplicative, and repetitive, and are likely to be less cooperative in responding to a poorly constructed form.

The order of the questions should consider, first, the respondent; then the interviewer (if any); then the processing personnel; and, lastly, the analyst.

Sequencing questions in favor of the respondents tends to improve the quality of their answers. The least sensitive, most general, and simplest questions should be placed first. Beginning the questionnaire with a few non-threatening or easy-to-answer items tends to promote a more positive attitude on the part of the respondent. Moreover, if at all possible, demographic questions should not be located at the beginning of the questionnaire since some respondents may find them threatening, e.g., questions about age, income, employment status. Usually it is best to place them close to the end, so refusals won't affect answers to earlier questions.

- Explanatory and Control Information

In addition to the questions themselves, survey questionnaires contain a variety of explanatory and control items to guide people who will be handling the forms -- respondents, interviewers, and data processing personnel. Don't neglect these items in your review.

Below are suggestions for critiquing the following "special" questionnaire items: (a) introductory explanations to respondents or interviewers; (b) instructions to whoever completes the questionnaire; (c) definitions; (d) interviewing aids such as show cards, calendars and scales, which interviewers sometimes use to prompt replies; (e) control numbers to identify the questionnaires and control their flow through the collection and processing operations; and (f) codes and directives for processing personnel.

= Introductory explanations.

Virtually all questionnaires contain a few explanatory remarks at the beginning of the questionnaire, either for the respondent or to suggest the interviewer's opening remarks.

Introductory information on a mail questionnaire is very important because no interviewer will be present to tell respondents (a) what the study is about, (b) its objectives, (c) why their cooperation is important, (d) how their replies will be used and who will have access to them, and, (e) how to get help if they have any problems.

Respondents also should be told at the outset that accurate and complete answers are desired and that they should think carefully, search their memory, and, if appropriate, take time to check their records. If any questions are particularly sensitive or threatening, a few additional comments may be necessary.

Introductory information should be included in a one-page letter accompanying the questionnaire. The letter should be individually addressed, if possible. (The mail merge capability of most word processors makes this feasible at little extra cost.)

A mail questionnaire also should advise respondents what to do with the questionnaire when they have completed it. Should the questionnaire be returned in a self-addressed envelope? What's the deadline for completing it? (Note that deadlines will increase the response rate.) A return address should appear on both the cover letter and the questionnaire proper.

Suggestions for the interviewer's opening remarks are usually stated at the top of the questionnaire. These should be brief. Long explanations tend to make respondents uncomfortable. The interviewers should simply identify themselves and the organization they represent, and state the purposes of the survey in one or two sentences.

= Instructions.

Instructions to respondents or interviewers on how to complete the questionnaire must be carefully phrased to prevent errors and omissions. Review the instructions as attentively as you do the questions.

All instructions should be uniform in style and clearly distinguishable from other material on the questionnaire, e.g., set off in capital letters. Only instructions applicable to all interviewing situations normally should appear on the questionnaire. (See "Item Placement" later in this section for additional review considerations.)

There are two basic kinds of instructions:

(1) Directions on how to answer the individual questions.

(2) So-called "skip instructions," which instruct the person completing the form where to go next, depending on how they answer the current question.

Skip instructions should be (a) worded positively and (b) reference a later question. They should inform the person completing the form where to skip to when a particular reply is given, *not* where to go when no answer is given. Skip instructions should never ask the respondent to skip backwards to a previous question.

Complex skip patterns should be avoided, especially on mail questionnaires. (They are easily managed in a computer-assisted telephone interview, however, because the system can be programmed to present the next question correctly, based on the last answer keyed in by the interviewer.)

Note that, in addition to the instructions printed on the questionnaire, interviewers are given *separate* question-by-question written instructions. These are usually more detailed and cover unusual interviewing situations. Usually they are incorporated in a manual and used both for training and reference purposes.

= Definitions.

In the interest of clarity, any unusual terms on the questionnaire should be defined. For example, if manufacturers are asked to estimate the "value of goods" sold last year, the questionnaire should indicate whether answers should be expressed in current dollars, the depreciated book value, or some other method of calculating the value.

Definitions of technical terms often are a major component of questionnaires for Agency-sponsored surveys. It is not unusual for an entire section to be devoted to definitions. Be sure to have the project personnel most knowledgeable about the subject matter review all definitions.

= Interviewing aids.

Although the visual aids that interviewers show respondents to encourage more accurate replies are not strictly a component of the questionnaire, you should review them along with the questionnaire to make sure they contain an appropriate range of alternative answers.

= ID and control information.

Every questionnaire should contain information to identify it and control its flow through the collection and processing stages. At a minimum, the first page or cover page should include the title of the study, the name of the organization conducting the study, the OMB control number and expiration date, and a space to insert whatever multi-digit code numbers the contractor plans to use to identify the response units for follow-up, evaluation, or cross-referencing purposes and for determining what sample weights to apply (see Chapter 4). (Since it is possible for the questionnaire to come apart, each page should be numbered and include some information identifying the form.)

In addition, in face-to-face or telephone surveys, there should be a space to record the date and time the interview began and ended. The contractor also may include a place to rate the performance of the interviewer or processors.

Make sure that proper identification and control information is included on the final draft of the questionnaire. Check these items again when you review proofs of the final questionnaire.

= Data processing provisions.

If at all possible, the format of the questionnaire should be arranged so it is easy for the transcribers or the data entry clerks to proceed from one item to the next. Certain formats and coding schemes can simplify the processing operations and, at the same time, facilitate the tasks of the respondents or the interviewers.

Closed questions can be "precoded" to facilitate processing and ensure that the data are in proper form for analysis. Precoding involves assigning a code number to every response option. The response options are either explicitly stated in the question or are printed on a card handed to the respondent. When they appear on the questionnaire, the respondents select their replies by checking a box, circling a coded answer, underlining a preprinted response option, or writing in a code or a number. Provisions also may be made for "No answer" or "Don't know" replies.

When the completed questionnaires are processed, the data entry clerks simply key the appropriate numerical codes directly into the computer. This eliminates one step in the processing because the replies do not have to be coded or transcribed onto a coding or keying sheet before being entered into the computer.

3. Reviewing the Format

The last step in your review should be devoted to the the general format of the questionnaire, specifically to the --

- General appearance
- Length
- Placement of questions and instructions

GENERAL FORMAT

Although the contractor should have designers experienced in the proper formatting of questionnaires, a final review by Agency subject-matter and data processing specialists may suggest revisions that will improve the questionnaire's response-getting power.

A well-formatted questionnaire can significantly reduce response errors. If the questionnaire is designed to be self-administered, your review of the format should have high priority. The format should give primary consideration to the respondents, then the interviewers, and lastly the data processors.

- General Appearance

The general appearance of the questionnaire, the kind of paper it is printed on, the size and style of the type, and the amount of open space all influence how well the respondents or the interviewers are able to follow instructions and complete the questionnaire. Appearance is very important in a self-administered questionnaire and will influence the response rate.

The questionnaire form should look professionally designed and easy to answer. If the form is more than four pages long, a booklet format is desirable. It should be printed on good stock because it will be subjected to a great deal of handling during the course of the collection and processing operations.

Colored paper or color-shaded sections may be helpful in a complex questionnaire. Shading can be used to direct attention to answer spaces, to highlight certain topics, to indicate transitions between sections, and to reserve space for office use. The reduction in respondent and clerical errors is well worth the small additional expense for two-color printing.

Large, clear type should be used throughout. Different type styles should be used for questions, instructions, and data processing notations. Instructions should be in bold type so they are clearly distinguishable from the questions.

Above all, the questionnaire should not look crowded. Ample white space should be allowed because it will make the questionnaire look easier to complete, and generally will result in fewer errors by both interviewers and respondents. Response formats should be consistent, and adequate space should be allowed for replies to open questions, arithmetical calculations, and general remarks (by respondents or interviewers).

- Length

Survey literature abounds with recommendations on questionnaire length. The general consensus is that setting an arbitrary limit on length is unnecessary and unrealistic. Much depends on the method of administration, the respondent's obligation to reply, the subject matter, and the way the questionnaire is constructed.

Let's first look at the length of self-administered questionnaires. Since no social interaction is involved, mail questionnaires sent out to the general public are directly affected by length. If the subject matter is interesting and relevant and the respondents are generally well-educated, the questionnaire may be 12-16 pages long and there will be no serious loss of cooperation. If the topics are likely to be of little interest to the respondents, however, the questionnaire should not exceed four pages. Anything longer is likely to induce fatigue and result in a considerable number of response errors and a lower completion rate. Even poorer response can be expected if efforts to cut down on length include crowding questions, using oversize paper, or reducing the print size.

The length of a mail questionnaire is not as important in a business survey. In fact, EPA relies heavily on long, complex, self-administered questionnaires for obtaining detailed technical information from business and industry. Whether replies are voluntary or mandatory, a long mail questionnaire is often less burdensome than a lengthy face-to-face interview. It is less disruptive of office routines and gives the organizations time to discuss the questions with other people and search their records, as necessary.

As for interview surveys, if the topics are interesting and important to the respondents, face-to-face interviews of two or three hours can be conducted with little difficulty, regardless of the type of respondent. Telephone interviews lasting over an hour also can be conducted successfully provided they deal with highly salient topics. On the other hand, unless responses are mandatory, telephone or face-to-face interviews have to be considerably shorter -- 20 to 45 minutes at most, as a rule.

Remember that the length of the data collection instrument directly affects the total response burden of the survey. "Response burden" is the time it takes to complete the data collection instrument.

The estimated amount of time it takes to complete the proposed questionnaire multiplied by the number of respondents in the sample is the total response burden you must report to OMB in your clearance request. The burden should not exceed the allowance provided for the survey in your office's Information Collection Budget.

- Item Placement

  The placement of the questions, instructions, and other items on the questionnaire can make the task of respondents and interviewers easier and more enjoyable. The placement of response categories also should be consistent. In some cases, good placement helps to minimize response errors, refusals, and incompletions.

  Below we discuss some general rules for the placement of (a) questions and (b) instructions. Placement "rules" for other items (introductory material, definitions, and ID and control information) were covered earlier in this section.

  = Questions.

  The questionnaire should start with a few short items that are relevant, interesting, nonthreatening, and necessary. As we mentioned earlier, placing questions the respondent may perceive as threatening at the beginning of the questionnaire may result in defensive -- and frequently invalid -- responses. It is best to put them close to the end but not at the end of the questionnaire. Important questions should be placed towards the beginning. The last items in a questionnaire rarely get the same degree of attention as earlier ones, hence the least significant items should be placed last.

  It is generally best to start a mail questionnaire with a few short, simple, closed questions. Never begin with an open question requiring a lengthy response. Writing long answers may be difficult and embarrassing for some respondents, who may worry about making spelling and grammatical errors. Also, include space at the end for general comments.

  Questions should never be split between two pages because the person completing the form may think the question is complete and inadvertently provide a premature, inaccurate response.

  = Instructions.

  Instructions on how to answer a question or a series of questions should be placed before items, not at the beginning of the questionnaire.

Instructions for responding to individual items should be placed either immediately before the question or immediately after it, prior to the space provided for the answer.

Skip instructions should be placed immediately after the answer space allowed for the question. Sometimes words and sometimes arrows are used to advise respondents or interviewers which question they should answer or ask next, depending on how the current question was answered.

Coding or probing instructions for interviewers should be placed after the question. Notations for coding personnel should be in small type and located so they will be as unobtrusive as possible to respondents or interviewers.

C. MONITORING PRETESTS

In addition to reviewing all questionnaire drafts, the project officer should take an active role in both the exploratory research and testing activities. Time spent in testing the questionnaire before collecting data for the main survey may eliminate problems that would be costly if not impossible to correct later. For this reason, a pretest and a pilot test are vital for a major survey.

Before the contractor is hired, be sure to review the pretest provisions of the offerors' proposals. (See section B-3 of Chapter 6, Volume I, for more information.)

After the contractor is aboard --

(1) Participate as an observer in any exploratory interviews that many be conducted. This will help you evaluate response problems, some of which may be serious enough to require changes in the data requirements of the survey.

(2) Critically review the contractor's plans for testing the questionnaire. Make sure that (a) the pretest sample adequately represents all important subgroups of the target population, including any types of respondents for which special problems are anticipated; (b) the size of the test sample is adequate for a valid test; (c) the test conditions approximate those of the actual survey; and (d) enough time has been allowed to analyze the test results and incorporate any necessary revisions in the questionnaire before the survey starts.

(3) Facilitate and coordinate all internal approvals required by your office for pretest(s) as well as the OMB clearance request (if more than nine respondents will be used) with the EPA's Statistical Policy Branch and Information Management Branch of the Office of Standards and Regulations and your office's information management coordinator.

Clearance requests for pretests may be submitted separately or in combination with the clearance request you submit for the main survey. (See section B of Chapter 7, Volume I, for details.)

(4) Participate in all pretests. Go along on a few of the interviews to get a first-hand view of respondents' reactions to the questions, and attend the debriefing sessions at the conclusion of the tests.

(5) Review all pretest reports carefully. They should include a list of the proposed refinements to the questionnaire and an analysis of the pretest data. The pilot test report also should propose any refinements to the field procedures the contractor deems necessary.

When you review subsequent drafts of the questionnaire or plans for further tests, make sure the contractor has taken into account all reviewers' suggestions.

FOR MORE INFORMATION ON
QUESTIONNAIRE DEVELOPMENT --

- Approaches to Developing Questionnaires, Statistical Policy Working Paper 10, Statistical Policy Office, Office of Information and Regulatory Affairs, OMB, Washington, DC, 1983.
- Asking Questions: A Practical Guide to Questionnaire Design, S. Sudman and N. Bradburn, Jossey-Bass, San Francisco, CA, 1982.
- "Questionnaire Construction and Interviewing Procedures," Research Methodology in Social Relations, Fourth Edition, A. Kornhauser, P. Sheatsley, et al; Holt, Rinehart, and Winston, New York, NY, 1981.
- The Art of Asking Questions, S. Payne, Princeton University Press, Princeton, NJ, 1951.

SOURCES OF QUESTIONS FOR
HOUSEHOLD SURVEYS --

- Basic Background Items for U.S. Household Surveys, R. Van Dusen and N. Zill, Social Science Research Council, Washington, DC, 1975.
- General Social Surveys, 1972-1982: Cumulative Codebook, National Opinion Research Center, University of Chicago, Chicago, IL, 1982.
- Measures of Social Psychological Attitudes, Revised Edition, J. Robinson and P. Shaver, Institute for Social Research, University of Michigan, Ann Arbor, MI, 1973.

SAMPLING

Sampling is selecting some portion of a target population, sometimes called a study population or simply population, and investigating just this portion, which is called a sample.

A half century ago, many statisticians felt that collecting information about every member of a population they wanted to investigate was the only acceptable way of conducting a survey. Today, as a result of technical advances in sampling theory and its applications, sample surveys are now widely accepted as an efficient and reliable way of studying individuals, land areas, or even extremely unstable environmental media such as surface water or air. Thus, specimens of blood or urine constitute a sample of a patient's body fluids. Specimens of soil taken from a lawn comprise a sample from that lawn. Specimens of water from a swimming pool form a sample of the water in that swimming pool. And so forth.

In this chapter we'll give you an overview of the basic concepts of sampling theory and some practical tips on monitoring the sampling activities of a survey contractor. We shall consider two general types of sampling: Probability sampling, which refers to the selection of sample members by chance, and non-probability sampling, where the units selected for study are chosen according to some purposive or convenient scheme, often by expert judgment. Specifically, we'll look at --

- The advantages of using sampling for an Agency-sponsored survey;
- The relationship between sampling errors and sample size;
- The methods used to design survey samples;
- The major components of a sampling plan; and
- Ways the sponsoring office can ensure the quality of the sampling activities.

A. ADVANTAGES OF USING SAMPLING

Almost all statistical surveys the Agency sponsors use sampling to select the members of the population they want to

Preceding page blank

study. Why collect information from only a sample rather than everyone in the population?

In most research situations, taking a census of the study population is both impractical and inefficient. The most important reason for investigating only a sample of the population generally is to hold down costs. Obviously it is cheaper to collect information about 500 people, land areas, processes, etc., than about 5,000, say. Fewer staff are needed to collect the information and process it in a form suitable for analysis. Using sampling for studies of human populations also reduces the burden on those from whom information must be collected. Sampling also gives faster and more accurate results because fewer data have to be collected and processed.

Let's expand on these four main advantages of sampling.

1. Lower Costs

   If the population of the proposed study is very large -- national in scope, say -- collecting information about the entire population is simply out of the question from a cost standpoint. The cost of taking a census of the U.S. population in 1980 was $1 billion, for example. A good quality sample survey of a large human population requires a small fraction of the resources needed to collect data from everyone in the study area. The per-unit cost of a sample normally is higher than complete enumeration of the population because more highly trained staff and more stringent quality control throughout every phase of the survey are required.

   Similarly, if you plan to use an expensive measurement procedure to collect certain environmental data, studying a sample of the population often is the only feasible way of keeping costs within reasonable bounds. The cost of using an expensive monitoring device to measure ambient air quality in more than a small number of communities may be prohibitively expensive -- as well as unnecessary, given the state of the art of sampling.

2. Reduced Paperwork Demands

   The Office of Management and Budget, in accordance with the Paperwork Reduction Act of 1980, imposes limits on all Federally-sponsored information collections. Using sampling to study a population of interest helps to minimize the paperwork demands Federal agencies impose on the public, particularly on business and industry.

3. More Timely Results

The Agency often needs the results of their survey research projects quickly. Because fewer respondents or specimens have to be investigated in a sample survey, the time required to collect and process the data generally is substantially lower.

4. More Accurate Results

Since survey researchers use carefully controlled procedures to collect and process sample data, it is not unreasonable or unlikely for a well chosen sample to produce more accurate results. Although sampling introduces a source of error in the data -- called sampling error -- that would not occur if all members of the population were studied, sampling error is identifiable and measurable.

At the same time, because the investigators focus the available resources only on a portion of the population, there is less chance for human error and, therefore, the data quality tends to be higher. Human errors can creep in at any stage of a survey -- during the data collection phase, during the editing and coding of the questionnaires, and during the tabulation and analysis operations. Because there are fewer data to deal with in a sample survey, greater quality-control can be exercised throughout each stage to guard against all kinds of errors.

Given these advantages, are there any research situations where sampling may not be appropriate for collecting environmental and health data which EPA needs to effectively fulfill its Mission?

In some cases, only sampling is possible -- air or water monitoring, for example. In studies of human populations, if the study population is small or if separate detailed data for small subsets of the population are desired, collecting data for the entire population may be appropriate -- at least for some parts of the investigation. If your target population is all U.S. chemical manufacturers, for instance, it probably would be feasible to study only a sample of them to get the information you need. However, if you were interested in a specific chemical produced at only ten plants in the United States, it probably would be best to collect data from all these plants. Similarly, if you were interested in all the chemical manufacturing plants in a single county, it might be best to survey all of them.

B. SAMPLING ERRORS AND SAMPLE SIZE

In Volume I we recommended that, in establishing the minimum design criteria for your survey, you include an acceptable level of sampling error for the key statistics you need to achieve your research objectives. Since this is a task that should be done in the planning stage, before a contractor is hired, we'll discuss sampling errors before considering other aspects of sampling. We'll also show you how sampling errors are measured, and the relationship between sampling errors and sample size.

The purpose of most surveys is to measure certain characteristics of a population. When only a portion of a population is used for study purposes, survey statisticians need a way of estimating the extent to which this portion -- the sample -- and the entire population differ from each other. Studying a sample rather than every member of a population means abandoning mathematical certainty and entering the realm of inference and probability. The values of the estimates (statistics) derived from the data collected from the sample, by the same token, also will be different than the actual mathematical values that would have resulted had data been collected for everyone in the population. The difference in these two sets of values for every statistic is called the sampling error. Collectively, sampling errors are errors that statisticians can measure and take into account in reporting the survey findings. Other sources of data errors in a survey are (a) estimation biases, (b) systematic errors caused by defective measuring devices, (c) exclusion of part of the population due to a faulty sampling frame, and (d) failure of the interviewers to ask all the questions -- all of which produce errors that are much more difficult to measure than sampling errors and which can significantly affect the survey results.

1. Sampling Errors

Sampling errors, we have seen, are measures of the extent to which the values estimated for the sample such as means, totals, or proportions differ from the values that would be obtained if the entire population were surveyed. Since there are inherent differences among the members of any population, and since data are not collected for the whole population, we cannot know the exact values of these differences for a particular sample. Moreover, different samples give different results. To compute sampling errors, therefore, statisticians measure the average differences between sample estimates and population values, i.e., averages of the differences for a hypothetical set of sample surveys using the same sample design and measurement procedures.

When a probability sample is used, sampling errors can be estimated with a certain degree of precision. A probability sample is one in which each member of the target population has a known, positive probability of being selected. In fact, the main reason we have recommended that probability sampling be used for all Agency surveys, whenever feasible, is that statements based on sample results are always probability statements -- always estimates, not statements of fact. If probability methods are not used to select the survey sample, there is no way of knowing how much error there is in the data and hence how much confidence one can place in the survey findings.

2. Measuring and Expressing Sampling Errors

Let's look now at the ways statisticians measure and report sampling errors when probability methods have been used to select the study population.

Suppose you have contracted for a survey to determine how many families in a particular city -- we'll call it City X -- are getting their drinking water from contaminated sources. Now, after completing the survey, let's say the contractor estimates that 40 percent of all families in City X are using contaminated sources. The contractor tells you that the standard error, or standard deviation, of this estimate is 2 percentage points. Moreover, the contractor says that this estimate is likely to be within 4 percentage points of the true proportion of families in City X using contaminated water. What does this mean?

The standard error is a measure of the probable accuracy or precision of any one estimate derived from sample data. To relate the standard error of this particular statistic -- that 40 percent of all families in City X are using contaminated sources -- to the true value, the contractor formed a 95 percent confidence interval, which is approximately defined as --

Sample estimate ± twice the standard error (S.E.)

The confidence interval in this example is the interval from 36 to 44 percent, i.e., 40 percent ± 2 x 2 percentage points.

Provided the contractor has used a reasonably large sample of families in City X to collect data on the quality of the drinking water, you could give 19 to 1 odds that this 95 percent confidence interval would include the value you would get if you surveyed all the families in City X. If you were willing to accept lower odds or if you wanted higher odds, other multiples of the standard deviation could be used to attain other confidence levels, such as --

Confidence Interval	Approximate Level of Confidence
Estimate ± (1.0 x S.E.)	68%
Estimate ± (1.6 x S.E.)	90%
Estimate ± (2.0 x S.E.)	95%
Estimate ± (2.6 x S.E.)	99%

Let's turn to another aspect of reporting sampling errors. Sampling errors may be expressed either in absolute or relative terms. To illustrate the difference, let's suppose that City X has a total of 5,000 families. The 40 percent estimate of families using contaminated drinking water translates to a total of 2,000 families.

Exhibit 4 on the next page shows the absolute and relative sampling error of this estimate expressed in three ways. Relative standard error (relative to the estimate) is often called the coefficient of variation. It is always expressed as a percentage of the sample estimate. As you can see, the relative standard error (or the coefficient of variation) is the same for each type of estimate, even though the estimates themselves and their standard errors are expressed in different units.

When you establish the Agency's minimum design specifications, therefore, be sure to state whether you are referring to absolute or relative sampling errors. This is especially important for estimates of percents or proportions.

3. Determining Sample Size

How large a sample is needed for a particular survey? Questions about sample size seem to be simple ones, but answering them is not so simple.

In Chapter 3 of Volume I, we recommended that you exclude the size of the sample when you specify the

EXHIBIT 4

ABSOLUTE AND RELATIVE SAMPLING ERRORS FOR DIFFERENT TYPES OF ESTIMATES OF FAMILIES USING CONTAMINATED DRINKING WATER SOURCES

Type of Estimate	Sample Estimate	Standard Error	Coefficient of Variation
Total	2,000 families	100 families	5%
Proportion	0.40 of families	0.02 of families	5%
Percent	40% of families	2% of families	5%

survey design criteria in the statement of work for the RFP. The level of sampling error (or level of precision, as it is sometimes called) and sample size are closely related. When probability sampling is used, it is relatively easy to determine how many members of the target population have to be included in the sample to achieve results with the level of precision you have specified. For a particular sample design, it is primarily the number of units in the sample not the percentage of the population the sample represents that affects the precision of the sample estimates. For example, in estimating percents or proportions, the sampling error associated with a sample of 1,000 units taken from a population of 100,000 is almost the same as the error for a sample of the same size from a population of 100,000,000.

We recommend that you specify the level of precision you need for the key estimates (statistics) and leave it to the offerors to propose a sample design that meets this specification at the lowest possible cost. If you specify both precision and sample size, the offerors may find it impossible to meet both your requirements.

To achieve the most efficient sample design, the contractor must determine a sample size that --

(1) Will achieve a fixed level of precision for minimum cost; or

(2) For a fixed cost, will achieve the greatest estimation precision.

In virtually all EPA survey contracts (1) will apply. In other words, the contractor starts with a requirement to attain a given level of accuracy (precision) and must satisfy this requirement at minimum cost. Alternatively, the contractor may be given a particular budget and must make a sample allocation that will provide the most accurate results. By "allocating the sample," we mean dividing the sampling units (the units of the population from which the sample is drawn) among various components of the sample such as strata, regions, counties, cities, and so on.

How many sample members are taken from where? An example of the difficulty a contractor may encounter in allocating a sample in a study of environmental media is the following: If we have the capacity to chemically analyze 1,000 specimens of lake water, how many sample lakes and how many specimens per lake are most efficient in answering our questions?

When you draft the statement of work for the survey, be sure to consult a sampling specialist to ensure that the precision levels you set are reasonable given the resources you have available.

In addition to the levels of precision you specify in the statement of work for the key statistics, the offerors also have to take the following design factors into account in determining the sample size.

- The level of geographic detail for which estimates are needed. If the target population is the entire U.S. population, getting estimates at a specified level of precision for each State would require a sample roughly 50 times larger than that required to get estimates with the same level of precision for all 50 States collectively.

- Variability of the characteristics of the target population. The greater the differences between the units in the target population, the larger the sample has to be to achieve a specified level of precision. The level of precision in sample surveys, in fact, is based on sample variance. It measures the lack of homogeneity among the data collected from the sample.

- The methods used to design the sample. Survey designers use many different sampling methods and combinations of methods to design a survey sample. The levels of precision for a sample of a given size will vary, depending on the sample design.

Cluster sampling,a method of choosing a survey sample in which all the sampling units are clustered in one or more geographic areas rather than across the entire area in which the population is located, has perhaps the greatest impact on the precision of the statistics. (See section C below for more about cluster sampling.) Obviously, estimates derived from a sample of 1,000 households chosen at random from a city directory would give a considerably higher level of precision than those derived from a sample of only 50 households chosen from each of 20 randomly-selected city blocks.

- Expected level of non-response. In almost all sample surveys, regardless of what method of collection is used, researchers will not succeed in obtaining responses for every unit in the sample. There are many reasons for this, which we'll discuss in Chapter 5. For example, a respondent may refuse to be interviewed, or an interviewer may fail to contact an acceptable respondent, or the person designing the sample may include ineligible units (such as a business that is no longer active) in the sampling frame. The sampling frame is the list of units from which the sample is drawn.

  Often, survey designers increase the sample size to compensate for the anticipated rate of non-response. This will reduce sampling errors, but it will not reduce the bias in the estimates that arises because eligible units provide no data or incomplete data.

- Cost and time. As we indicated above, the resources the Agency has available to do the survey place constraints on the size of the sample -- generally, the larger the sample, the more the survey will cost.

  Moreover, if there is a deadline for obtaining the results, the time it will take to collect and process the sample data also may limit the size of the sample.

C. SAMPLING METHODS

In this section we will describe briefly the methods most commonly used to design survey samples. To illustrate the different methods, we will continue with the City X example used in section B. Knowing something about the different methods used to construct a sample will give you a better understanding of sample designs you may have to review.

Our focus in this section is on probability sampling methods, which we recommend for virtually all Agency surveys. Probability sampling, also called random sampling, is an objective process recognized and accepted as standard procedure by knowledgeable survey specialists throughout the world. We will also describe three types of non-probability samples. Non-probability samples are selected according to some purposive or convenience method, often by an expert or specialist on the basis of his or her considered opinion.

1. Probability Sampling Methods

Probability samples are those in which the members of the population (or the sampling units) are selected at random -- solely by chance. "Random" is not equivalent to "haphazard." A true random selection must be independent of human judgment. The two distinctive features of probability sampling are --

(1) The use of some random device (such as a table of random numbers) to determine which units in the population (or the frame) are included in the sample. This prevents the person designing the sample from biasing the selection (consciously or unconsciously) towards a sample that will produce some desired result.

(2) The sample can be used to make estimates of the sampling errors associated with the survey findings. Hence, anyone using the survey data can determine how accurate the data are and how much confidence to place in any conclusions based on the sample data.

Let's look at six of the most common methods of probability sampling used today:

- Simple random sampling
- Stratified sampling
- Cluster sampling
- Systematic sampling
- Sampling with probability proportionate to size
- Multi-stage sampling

PROBABILITY
SAMPLING
METHODS

- Simple Random Sampling

In simple random sampling each unit in the target population has an equal chance of being selected, a characteristic shared by many probability sampling methods. Simple random sampling is also known as "sampling with equal probabilities," or "equal probability selection." However, simple random sampling is unique in that every possible sample of a given size has the same probability of being selected.

Simple random sampling is particularly appropriate for very small studies where the sampling units are approximately the same size or there is no useful measure of size for the survey. A sample of medical records in a hospital (to review diagnoses for possible cases of pesticide poisoning, say) is an example of a situation where simple random sampling may be appropriate. Simple random sampling is seldom used by itself in designing Agency surveys, but it is frequently used in combination with one or more of the other sampling methods described in this section.

Let's see how we would draw a simple random sample from the 5,000 families in City X. First, we would need to prepare a list of all 5,000 families. We might get this from a city telephone directory, or it may be necessary to create a list by canvassing the area or some other means. We would then list all the families by name and number them in sequence from 1 to 5,000.

To begin the selection of the sample, we would pick a random number between 1 and 5,000 -- 254 say. The family with that number would be the first unit included in the sample. We would continue to randomly select numbers until we had chosen the desired number of sample units -- 500 families, perhaps.

What if the same random number comes up more than once? Usually, numbers that have already been picked are set aside so that no number (the number "254," for example) shows up more than once. This is known as simple random sampling without replacement, i.e., a number, once selected, is not returned to the sampling frame. (Note that sampling with replacement, where the numbers are returned to the frame, is sometimes used for probability samples, including simple random sampling.)

- Stratified Sampling

It is often useful to divide the population into exhaustive and mutually exclusive subgroups for sampling purposes. If we propose to sample from every subgroup, then the subgroups are termed strata. In stratified sampling, the population (or the frame from which the sample is drawn, if they are not equivalent) is divided into two or more strata, and the selection of the sample is carried out separately for each subgroup or stratum. Stratification does not imply any departure from probability selection. It only means that before any units are selected, the population is divided into one or more strata. Then a random sample is selected within each stratum.

Continuing with our City X example, let's suppose we have reason to suspect that contamination is more likely to occur in some parts of City X than in others. If so, we could use a geographic stratification to select the survey sample. For example, we could draw a separate sample from each of the city's seven wards. This would ensure the selection of some sampling units in each ward, whereas if we did not stratify, the sample could -- purely by chance -- be heavily concentrated in one or two wards.

How should the overall sample be allocated among the strata, or wards? If we had no clue as to the likelihood of contamination in different strata, we would probably use the same sampling fraction, say, 1 in 10, in each of the wards. This is called proportional stratified sampling because the distribution of the sample families in each ward would be proportional to the distribution of families in each ward in the population.

It is not necessary to use the same sampling fraction in each stratum. If we had information indicating that the drinking water contamination problems were much more serious in three of the seven wards, we could sample at a higher rate in those three wards.

The primary reason for using stratified sampling is to make the sample more efficient, i.e., to produce estimates with smaller sampling errors. How well this objective is met depends on the criteria used to define the strata.

- Cluster Sampling

In cluster sampling, groups or "clusters" of adjacent units in the population are formed and a random sample of the clusters is selected. In other words, within a particular stratum, rather than selecting individual units one by one, whole clusters of units are selected.

To illustrate cluster sampling, one way of selecting a probability sample of families in City X would be first to select a sample of city blocks at random and then construct a sample of some or all of the families living in those blocks. If City X has a total of 100 blocks, we might use simple random sampling to choose 10 blocks and then interview some or all the families in only these 10 blocks.

Estimates derived from a cluster sample are likely to have considerably larger sampling errors than estimates from a simple random sample having the same number of families. The reason is that adjacent sampling units tend to have similar characteristics. This similarity, or correlation, reduces precision by producing a degree of redundancy in the data collected from members of the same cluster.

Why, then, should we use cluster sampling? It is a practical necessity to use clusters in large surveys. First, there is a considerable savings of time and expense in compiling a frame that lists only the units in the clusters rather than all the units in the population. Second, if face-to-face interviews will be used to collect the data, by concentrating them in a smaller geographic area, the overall cost savings can be enormous -- especially in a national sample.

- Systematic Sampling

In systematic sampling, researchers first list the sampling units (which may or may not be individual members of the population) in some specific order. Then, they select units for the sample by computing an appropriate sampling interval (I) and and taking every Ith unit in the sampling frame. The starting point is chosen at random from the first I unit. This is called the random start.

To select a systematic sample of 500 families in City X from the 5,000 families in the frame, we might use a sampling interval of 10 (5,000 divided by 500) and a random start between 1 and 10. If our random start were "7" for example, the families included in the sample would be those numbered 7, 17, 27, and so on, up to the family with the number 4,997.

Systematic sampling is widely used in survey research, especially in combination with other methods. It has two main advantages --

= Only one random number need be picked during the selection process, rather than one for each unit needed to complete the sample.

= If the sampling units are listed in some meaningful order -- by block in City X, say -- the effect of using systematic sampling is essentially the same as using stratified sampling, i.e., certain types of units are assured adequate representation in the sample.

Another version of systematic sampling is sampling based on the ending digits of identification numbers. In this method, the last digit of a set of serial numbers that constitute the sampling frame is chosen at random, and all the units in the frame with ID numbers ending in those digits are included in the sample.

For example, suppose we listed the social security number (SSN) of the head of each family in City X. We could select our 1-in-10 sample by including all families with SSNs ending in "4." This method would give us a sample of approximately 500 families, although the exact size would depend on which ending digit was chosen as the random start.

Caution must be used in selecting any series of ID numbers for sampling purposes. SSNs frequently are used for sampling based on ending digits. For business surveys, IRS employer identification numbers (EINs) may be appropriate; however, because of certain peculiarities in the way EINs were initially issued, they are not suitable for serial sampling until the ending digits are assigned a more nearly random distribution.

- Sampling with Probability Proportionate to Size

Up to now, all the methods we have looked at have involved sample designs where every member of the population, or the sampling frame, or at least the stratum has an equal chance of being chosen as part of the sample. However, in some sample designs, all the sampling units do not have the same selection probability. If the population characteristics in which the researchers are interested are related to the size of the sampling unit, and it is possible to obtain some measure of the size of the units, greater precision usually can be achieved by giving larger units a greater probability of selection. This is sampling with probability proportional to size (PPS).

For example, in sampling the U.S. population, researchers typically select Standard Metropolitan Sampling Areas (SMSAs), counties, or other sampling units with probability proportional to the number of individuals residing there. In a soil study, counties may be selected PPS with crop acreage as the size measure. Or for a study of rivers, hydrologic units may be selected with probability proportional to the miles of river they contain.

To illustrate, suppose we wanted to select a sample of 10 of the 100 blocks in City X. We could simply select 10 blocks with equal probability using either a simple random sample or a systematic sample. However, if we had a count of the number of families in each city block (from a recent census, a local telephone directory, or some other source), and the blocks varied quite a bit in size (number of families), a more efficient sample design might result if we gave the more populous blocks a greater chance of selection. (By "more efficient" sample design, we mean one in which the statistics will have a smaller margin of sampling errors.)

The selection procedure would be as follows:

(1) First, we would list all 100 blocks in some order and, alongside each block, list the count (the number of families residing there) and the cumulative total of these families, as in the table below.

(2) Then, we would divide the total number of families in City X (5,000) by the number of blocks

to be chosen -- 10 in this case. The result -- 500 -- is the sampling interval we would use for selection purposes.

(3) Next, we would select a random start number between one and the sampling interval. Let's use 213 for illustration purposes. We would then form a series of sample-selection numbers beginning with the random start and add the interval as many times as needed, i.e.,

213, 713, 1213, ... 4713

(4) Finally, for each sample selection number (213 or 713, say) we would choose the first block whose cumulative total equals or exceeds that number until 500 units are chosen for the sample. Note that each block ultimately will be represented in the sample. The table below shows how the first two blocks were selected, e.g., blocks 2 and 6.

Block No.	No. of Families in Block	Cumulative	Sample Selection No.
1	120	120	
2	220	340	213
3	50	390	
4	170	560	
5	90	650	
6	130	780	713
7	310	1090	

Sampling with probability proportionate to size (PPS) is especially applicable for selecting the first-stage units of a multi-stage design (discussed next). To use PPS sampling, it is necessary to have "measures of size" for all the units in the target population or frame, e.g., counts of families by block in City X. The measures of size need not be exact. It is sufficient for them to be reasonably close to, or correlated with, their actual sizes.

- Multi-Stage Sampling

Earlier we discussed a sampling method called "cluster sampling," where groups of units rather than individual units are used to form the sample. Multi-stage sampling refers to the process of selecting subgroups within the clusters chosen at a previous stage. All multi-stage designs are, in fact, cluster samples. For practical purposes, virtually all Agency-sponsored surveys use some form of multi-stage selection. Multi-stage designs are essential for any national survey, face-to-face survey, or survey using a widely dispersed sample.

Continuing with our City X example, suppose we did not have a current listing of the 5,000 families in the city. We might decide to use a multi-stage design to select our sample. Let's start by illustrating a two-stage sample design. In the first stage, we might select a sample of blocks using probability proportionate to size, as discussed above, based on approximate block counts from the best available source. Next we would prepare lists of all the families in the sample blocks. Then, by simple random sampling or systematic sampling, we would select a sample of families from the list of families residing in each of the blocks selected in the first stage. Briefly put, the sample design would be as follows:

= Stage 1: Selection of sample blocks.

= Stage 2: Selection of sample families within the sample blocks.

The most important advantages of multi-stage sampling are --

(1) Researchers can concentrate on a smaller number of areas, with a consequent reduction in time, staff, and dollars; and

(2) Researchers need only listings of the sampling units chosen at the previous stage, rather than a complete list of the population, e.g., in the above example, the families in the blocks selected in the first stage instead of a list of all 5,000 families in City X.

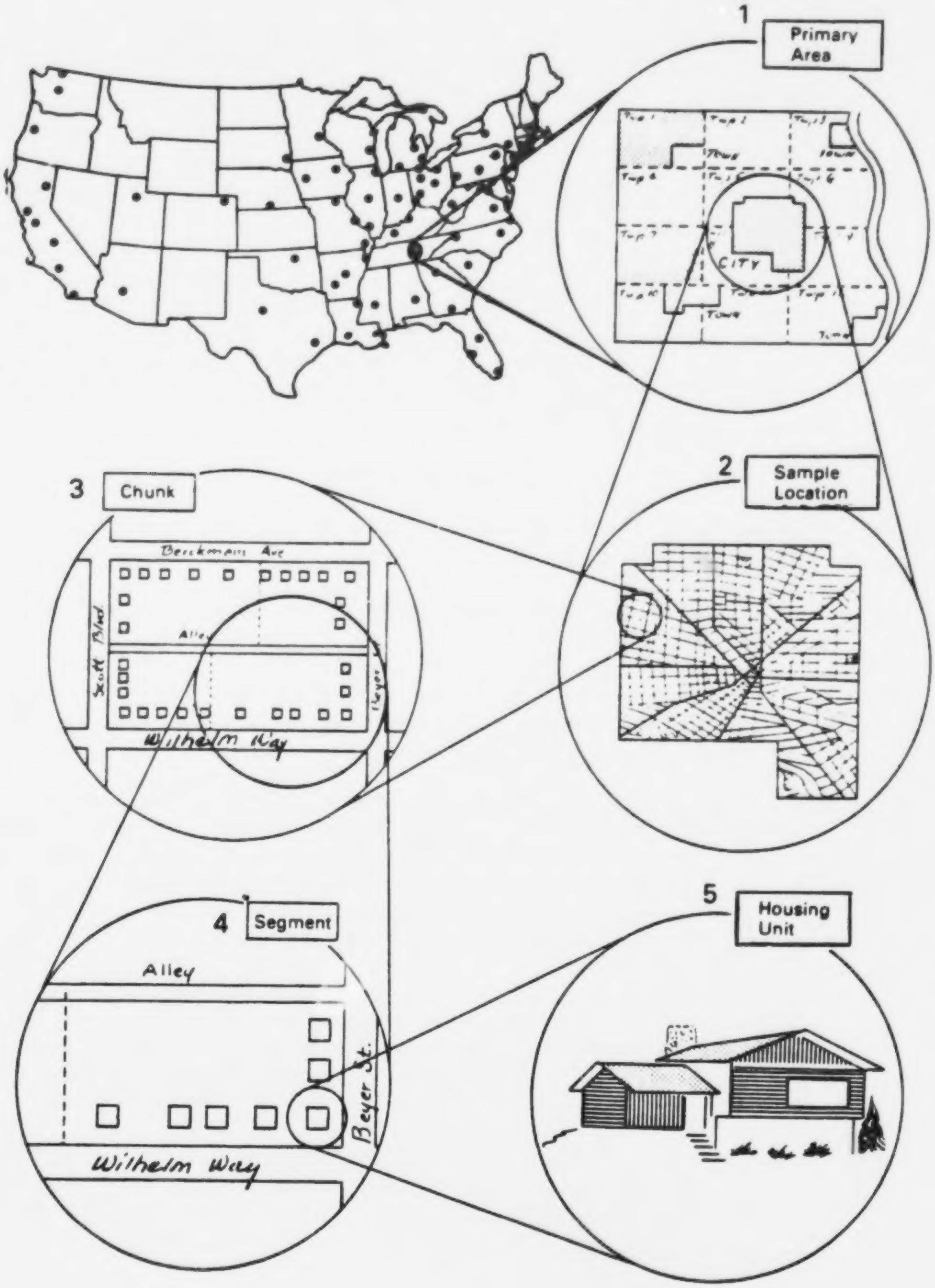
Most multi-stage samples involve four or five stages of selection. Exhibit 5 shows the stages of selection for a multi-stage household survey conducted by the University of Michigan's Survey Research Center. The stages of selection shown in the exhibit are --

= Stage 1: Selection of "primary areas," usually counties or groups of adjacent counties. In the Survey Research Center's design, 74 primary areas were selected (see any U.S. map).

EXHIBIT 5

MULTI-STAGE DESIGN FOR A NATIONAL HOUSEHOLD SURVEY

(Reproduced from "Interviewer's Manual,"
Survey Research Center, University of Michigan)


1 Primary Area
City
2 Sample Location
3 Chunk
Berckmans Ave
Alley
Scott Blvd
Wilhelm Way
4 Segment
Alley
Beyer St.
Wilhelm Way
5 Housing Unit

= Stage 2: Selection of "sample locations" (cities, towns, and rural areas) within primary areas.

= Stage 3: Selection of "chunks" (areas such as city blocks or rural townships, each containing from 16 to 40 housing units) from each sample location.

= Stage 4: Selection of "segments" of from 4 to 16 housing units in each sample chunk.

= Stage 5: Selection of "housing units" from the sample segments.

Our discussion of probability sampling methods has merely scratched the surface of the techniques survey statisticians use to construct samples and the ways they apply them to investigate various populations. Frequently, complex combinations of the methods we have described are used, along with variations such as double or sequential sampling, replicated sampling, and controlled selection.

There are several references at the end of this chapter that will help you expand your knowledge of probability sampling methods.

2. Non-Probability Sampling Methods

Non-probability sampling methods are characterized by a subjective selection procedure. Unlike probability sampling, the choice of the sample members is not random but, consciously or unconsciously, is influenced by human choice -- usually by expert judgment -- in accordance with some purposive or convenience scheme. The problem with all non-random selection schemes is that even the most conscientious individuals make unconscious errors of judgment that may be of considerable magnitude. These errors, which are very difficult to measure, are called "biases."

Because non-probability samples do have applications in some environmental research situations, we will briefly examine several types. Non-probability samples are also used sometimes in the final stage of selection of some environmental studies where strict probability sampling is not feasible, such as obtaining specimens for chemical analysis (house dust from a sample household, or water specimens from a small sample stream segment). They also are sometimes suitable for small-scale qualitative exploratory studies, and for pretests or pilot tests of EPA-sponsored surveys where the intent is to use probability methods to select the sample for the survey proper. Note that when non-random methods are used to select pretest or pilot test samples, the choice should not be restricted to "easy-to-get" units. If pretest samples include only units for which it is easy to collect information, it will be difficult to anticipate the kinds of problems that may occur in the main survey and how much the survey proper is likely to cost in time and dollars.

In any research situation where non-probability sampling is used, keep in mind that the results only pertain to the sample itself. The findings should not be used to make quantitative statements about any population, including the population from which the sample was selected.

Let's look now at the most common non-probability samples --

- Haphazard or convenience samples
- Purposive or judgment samples
- Quota samples

NON-
PROBABILITY
SAMPLES

- Haphazard or Convenience Samples

Haphazard or convenience samples are samples selected from populations for which it is relatively easy to collect information on a particular topic. Another feature of these samples is that the population groups from which they are selected do not reflect, with any measurable degree of error, the characteristics of some larger, well-defined group of which they are a part.

To illustrate, the following are examples of convenience samples of human populations --

= Voters interviewed in a shopping center;

= Volunteer subjects for experiments (e.g., families responding to a radio or newspaper appeal for volunteers to try out a new kind of water purification equipment in their homes);

= People answering a reader opinion questionnaire;

= People writing to their congressmen or senators about a particular issue.

- Purposive or Judgment Samples

  Purposive or judgment samples are samples that an investigator or another subject-matter expert considers to be "representative" of some study population. Like convenience samples, judgment samples are often used by EPA for pretesting purposes. For example, to pretest a survey of chemical plants that manufacture sulfuric acid, an expert researcher in the field might arbitrarily choose for preliminary investigation a few plants where all the manufacturing processes commonly used in the industry are represented.

  Judgment sampling is most usefully applied to early, exploratory phases of research involving extremely small samples. In environmental studies, judgment sampling and probability sampling are sometimes combined in a multi-stage sample, the final stage being a judgment sample.

- Quota Samples

  In some national surveys, investigators use probability sampling to choose the first one or two stages of a sample, and use quota sampling for subsequent stages. Quota sampling, therefore, is a version of stratified sampling in which the selection within strata is non-random.

  Quota samples also are frequently used in marketing and opinion research. For example, in an opinion survey, the interviewers will each be given a quota of interviews to conduct with various classes of individuals, households, businesses, etc. An interviewer's quota might consist of a specified number of individuals in each of six age-sex categories. Within these categories, and in the assigned area, the interviewer is free to decide how to locate and interview the specified number of individuals. However, since the selection process is subject to human judgment, there is no guarantee that biases will not occur. An interviewer may fill his or her quota in the top age group mainly with people 65 or 66, thus the very old will be underrepresented.

  Quota sampling has two main advantages:

(1) It is less costly than random sampling -- perhaps one-third as much; and

(2) There is no need to develop a frame for selecting respondents in the sampled area, which means that call-backs are avoided. If an eligible respondent is not available at a dwelling when the interviewer calls, the interviewer simply proceeds to the next dwelling.

As with all other non-probability samples, the non-randomness in the selection of the sampling units is the main disadvantage of quota sampling. Thus, it is impossible to estimate the sampling variability from the sample and to know the possible biases, which may be sizeable.

D. MAJOR COMPONENTS OF A SAMPLING PLAN

The starting point for developing a sampling plan is the five minimum survey design specifications we recommended for all Agency surveys in Chapter 3 of Volume I. These design specifications, which the sponsoring office should clearly define in the statement of work, are (a) the research objectives, (b) the target population and coverage, (c) the required level of precision (sampling error), and (d) the target response rate. The fifth design specification is that (e) probability sampling be used throughout the selection process whenever feasible.

In a contract survey, offerors normally will submit a draft of the sampling plan in their technical proposals. The plan may undergo several refinements before the final selection of the sample for the survey proper occurs.

The main components of a sampling plan, which are discussed in the remainder of this section are --

- Sampling frame(s)
- Sample selection procedures
- Estimation procedures
- Procedures for calculating sampling errors

COMPONENTS
OF A
SAMPLING PLAN

1. Sampling Frames

A *sampling frame* is a listing of population elements -- geographic areas, manufacturing plants, crop acreage, telephone numbers, city blocks, households, factories, etc.-- from which the survey sample is drawn. The frame is the most important component of the overall sample design because it identifies the population elements from which the sample is chosen. The population elements listed on the frame are called the sampling units. Often these are groups or clusters of units rather than individual units. The sampling units for which data are ultimately collected are known as the units of observation.

The choice of sampling frames and the steps taken to assure their completeness and accuracy affect every aspect of the sample design. Ideally, a sampling frame should --

- Fully cover the target population;
- Contain no duplication;
- Contain no "foreign" elements (elements that are not members of the population);
- Contain information for identifying and contacting the units selected for the sample; and
- Contain other information that will improve the efficiency of the sample design and the estimation procedures.

If the sample design calls for a multi-stage selection, a separate frame must be prepared for each stage (or stratum) of the sample design. For example --

- In the two-stage sample design for City X that we used earlier to illustrate multi-stage sampling, the frame for the first stage would be a listing of the blocks in City X. The frame for the second stage would be listings of all the families living in each sample block. In this two-stage design, the first-stage sampling units are the city blocks; the second-stage sampling units are the families for which the data will be collected. The families also are the units of observation.
- In a survey of plants manufacturing sulfuric acid, the sampling frame of the first stage might consist of a list of all U.S. chemical companies that manufacture sulfuric acid at one or more of their plants. After selecting a sample of these companies, we could make a listing of all or only a sample of the sulfuric acid plants belonging to the companies

chosen at the first stage. This listing would serve as the frame for the second stage of selection.

The development of the frame can be a major undertaking involving substantial effort and expense. Complete, current frames do not always exist. Many frames have missing units and some frames contain duplicate listings. Both of these frame imperfections cause biases if they are not detected before the selection is done.

To illustrate some of these points, a city telephone directory is a poor frame for a telephone survey of all local households. Studies show that as many as 20 percent of U.S. households have unlisted numbers or no telephones. Using the telephone directory, therefore, would result in undercoverage of the population. Moreover, some households would be overrepresented because they have more than one listed number. Finally, most directories also include business and other nonresidential numbers, some of which are hard to distinguish from residential numbers.

For surveys of businesses, it is especially difficult to obtain complete and current lists. Probably the best lists are those maintained for Federal programs like social security, income taxes, unemployment insurance, and the economic censuses. Unfortunately, these lists generally are not available to EPA and other Federal agencies, so other sources must be used -- commercial business lists or lists that EPA maintains of organizations that are required to comply with certain Agency regulatory requirements.

In general, perfect or ideal frames are seldom available. The sampling plan should always specify what steps the contractor will take to evaluate the frames and deal with any deficiencies such as missing or inaccurate elements.

2. Sample Selection Procedures

The sampling plan must provide complete specifications for the procedures to be used to select units from the frame at each stage of sampling.

Most sampling is done at a central location -- usually at the contractor's main office. However, for some of the later stages of sampling, the selection may be done in the field. For example, in a face-to-face survey the field supervisors may select sample housing units from block or segment listings prepared by the main office. Similarly, in a mail survey, if the contractor intends to conduct follow-up interviews with some of the people who do not send back questionnaires, procedures for selecting the follow-up sample should be described in the sampling plan.

The selection procedures in the sampling plan should specify --

- Any tasks necessary for reorganizing or otherwise refining the frame prior to selection, such as --
  - Screening to eliminate units that clearly are not in the target population; and
  - Transforming information about individual units into measures of size (necessary for sampling with probability proportionate to size).
- Whether the selection of sampling units (at each stage) will be with equal probability or with variable probability. If variable probability is to be used, the basis for assigning selection probabilities to individual units must be included.
- The sample sizes or intervals. If stratified sampling is used, sizes or intervals may vary by stratum. For some designs it may be necessary to obtain preliminary counts or other tabulations from the sampling frame to determine the most appropriate size or intervals.
- The specific probability mechanism to be used to select the individual sampling units or, for systematic sampling, the random starting point. If selection is manual, the use of random-number tables is recommended. If the selection is done by a computer, most systems will have access to a random-number generator.
- Any steps that will be taken to screen out ineligible sampling units, obtain better addresses, etc., after the initial selection is made.

3. Estimation Procedures

Estimation procedures are the methods used to convert sample data into estimates -- totals, means, proportions, and other statistics -- for the population. The actual preparation of the estimates (and the calculation of sampling errors, discussed below) is done towards the end of the data processing phase of a survey, but the procedures that will be used to obtain the estimates should be included in the sampling plan. The approach used for the estimations also plays a role in determining the size of the sample -- another reason for determining the estimation procedures early on. In addition, some kinds of estimates require the capture of certain data when the sample is selected, during the data collection phase, or during the processing phase of the survey.

The estimation procedures should specify how the contractor proposes to derive the most precise estimates possible from the sample data using statistical techniques such as (1) applying "weights" to give greater relative importance to some sampled elements than to others; (2) making adjustments to reduce the bias caused by eligible sampling units for which no data were collected; and (3) using auxiliary information obtained from the questionnaires, the sampling frames, or other sources such as administrative records, other surveys, etc.

We'll elaborate briefly on these three methods of enhancing data quality.

- Application of Weights

  When analyzing complex samples, statisticians assign weights (or multipliers) to adjust for (a) sampled elements for which the probability of selection was in some way unequal, (b) eligible units for which no data were collected (total non-response units), and (c) sampling units not included in the sampling frame (non-coverage errors).

  To explain --

  If all the sampled elements had the same probability of selection (sometimes called a self-weighting sample), survey analysts can obtain valid estimates of some statistics such as proportions, means, percents, and medians without weighting the data obtained from the sample. However, to estimate totals for the sample, all units must be weighted by the reciprocal of the sampling fraction. For example --

  If a simple random sample of 1 in 10 housing units has been selected, population totals could be estimated by applying a weight of 10 to the data for each housing unit sampled, or, similarly, by tabulating the sample data and multiplying the sample counts or aggregate by 10.

To illustrate how weights would be applied to adjust for unequal probabilities of selection, if a multistage sample were used and a sample of 10 city blocks were selected from a total of 50 blocks, and then every tenth family in these 10 blocks were selected for interviewing, the over-all selection probability for these families would be 1 in 50 --

$$\frac{10}{50} \times \frac{1}{10} = \frac{1}{50}$$

A uniform sampling weight of 50 would then be used to estimate totals from the sample data.

- *Adjusting for Missing Data*

The techniques that will be used to adjust for total non-response (eligible members of the sample that provide no data) are usually incorporated in the estimation procedures. The techniques used to make these kinds of adjustments are --

- *Reweighting* the sampled units by the inverse of the proportion of units that did respond. For example, if the proportion of the sample that responded was 0.80, a reweighting factor of 1.25 (1.00 divided by 0.80) would be used to adjust for the non-response. Reweighting factors often are computed separately by stratum or for each member chosen at the first-stage of selection. This allows for variations in the proportions of different categories or areas of the sample that responded.

- *Duplicating* the values reported by the sampled units to compensate for eligible units that did *not* respond. Information from *all* sampled units can be used in selecting the units that are duplicated. For example, the units to be duplicated could be selected from the same size or industry category or from the same geographic area as the non-responding units.

These kinds of non-response adjustments will reduce non-response biases but will not eliminate them entirely. The use of non-response adjustments is not an acceptable substitute for diligent efforts to collect data for all eligible units in the sample.

Note that other different techniques are used to adjust for missing data from single questionnaire items (called *item non-response*). (These adjustment techniques are discussed in Step 7 of Chapter 6.)

- *Using Auxiliary Information*

Survey analysts often can improve sample estimates by taking advantage of auxiliary information about the population, which may be taken directly from the sample (from the questionnaires, for example), from the sampling frames, or from independent sources. Auxiliary information is most often used to construct *ratio estimates*.

Suppose, for example, that we want to estimate the number of unemployed individuals in a national household survey. One way to do this is to tabulate the unemployed people in the sample and assign them appropriate weights based on their selection probabilities, a procedure known as *simple unbiased estimating*.

For example, suppose we have an estimate of the total population from an independent source at the time of the survey (the U.S. census, say). We could use this independent estimate to construct a ratio estimate of unemployed individuals as follows:

$$\text{Ratio estimate of unemployed individuals} = \text{Unbiased estimate of unemployed individuals} \times \frac{\text{Independent estimate of total population}}{\text{Unbiased estimate of total population}}$$

In other words, we would use the sample data to estimate the *proportion* of unemployed individuals and apply that figure to an independent estimate of the total population to derive a more precise estimate of the *number* of unemployed individuals in the population. If we had independent estimates of the population by age and sex, we could make separate ratio estimates of the number of unemployed individuals in each age-sex group and total them to get an estimate of the total number of unemployed individuals in the population.

Several different kinds of ratio estimation procedures are available, as are other procedures that make use of auxiliary information such as regression estimation. The choice of procedures will reflect the survey designer's judgment about how all relevant data from the sample itself, the sampling frames, and other sources can be used to develop the most precise survey estimates, i.e., how to make the best use of all available information.

In practice, weighting can be a complex task because a combination of adjustments is often necessary. Weights first may be assigned to adjust for unequal selection probabilities. These weights then may be revised to adjust for varying levels of response within the sample. Still further revisions may have to be made later to adjust the sample to known distributions in the population.

The sampling plan, therefore, should fully describe the estimation methods, formulas, or procedures the contractor plans to use to produce the survey estimates.

4. Calculation of Sampling Errors

Of all aspects of sampling, calculating (or estimating) sampling errors is the most technically complex. Most surveys collect data on a large set of variables and produce estimates for both the variables and their relationships to each other. It is impractical and usually impossible to calculate standard errors for all the estimates. Survey analysts, therefore, normally compute standard errors only for the key statistics and a few other selected estimates. From these calculations, they develop generalized models from which other standard errors can be inferred.

The sampling plan should specify --

- The estimates for which sampling errors will be calculated. (Standard errors should be computed for all key variables and a selection of other statistics.)

- The approach that will be used to calculate the sampling errors (formulas, methods, or software packages).

- Any assumptions or approximations implicit in the proposed approach.

The extent of sampling error depends on the design of the sample. The formula for calculating standard error found in most over-the-counter software packages is applicable only to simple random sampling with replacement designs. It will produce overestimates or, more often, underestimates of sampling error if applied indiscriminately to other sample designs.

The sample designs for most of the surveys EPA sponsors are complex, often involving a combination of multi-state and stratification sampling methods. For these complex designs, survey designers use a variety of approaches for calculating sampling errors such as the "Taylor expansion method," "balanced repeated replications," "jackknife repeated replications," etc. (See Kalton, 1983, for more information.)

In addition, several software packages have been developed recently for calculating sampling errors of estimates that are based on complex sample designs. The selection of suitable software poses difficulties because most packages treat the sampling units chosen at the first stage as being sampled with replacement, when, in fact, this is rarely the case.

(See Step 8 in Chapter 6 for more information on the application of these approaches to the calculation of sampling errors after the data are processed.)

E. MONITORING THE SAMPLING ACTIVITIES

The sponsoring office's greatest impact on the development and faithful execution of a sound sampling plan occurs in the design stage of the survey. Therefore, with the help of other Agency offices, we suggest that you, as project officer, do the following before the contract is awarded --

(1) Specify in the statement of work for the RFP the kinds of information offerors should include in their proposed sampling plans. The main components of a sampling plan -- the selection and development of the sampling frame, the sample selection procedures, estimation procedures, and the procedures for calculating sampling errors -- are discussed in section D of this chapter. (For further information on preparing the statement of work, see Chapter 5, Volume I.)

(2) Make sure the technical evaluation panel reviewing the responses to the RFP includes someone qualified

to evaluate the proposed sampling plans. Expertise in sampling theory and its applications to surveys is necessary to spot defects such as --

- Any (unnecessary) departures from probability sampling;
- Imprecise descriptions of the sample selection procedures;
- Sample sizes or sampling allocation rates that will not achieve the levels of precision you specified in the RFP;
- Incorrect estimation formulas or methods; and
- Inappropriate formulas or methods for calculating sampling errors.

(For additional information on what to look for in reviewing the sampling plan, see "Evaluating the Technical Proposals" in Chapter 6, Volume I.)

After EPA awards the contract, there are several things you can do to monitor the execution of the sampling plan.

(3) Sampling, perhaps more than any other aspect of survey methodology, is an area where expertise is vital for effective monitoring and control. Most statisticians are not experts in sampling theory. We recommend, therefore, that you have an expert in this special branch of statistics review the sampling plan before giving the contractor permission to proceed with the development of the frame, the selection of the sample, and other sampling operations. If a sampling expert is not available in your office, you may request assistance from the Statistical Policy Branch of the Chemicals and Statistical Policy Division within the Office of Standards and Regulations. Afterwards, make sure the contractor revises the sampling plan to incorporate any changes you or other review authorities suggest.

(4) Be sure the contractor tests the validity of the sampling frames before beginning the selection of the sample for the survey proper. Missing and duplicative sampling units can cause difficulty if they are not detected. Frame counts, broken down by geographic area and other characteristics, should be checked against information about the population that may be

available from other sources. The accuracy of totals for various kinds of industrial establishments may be cross-checked with the most recent economic census, for example. Sometimes, especially when using commercial business lists, it may be desirable to contact a small sample of the units in the frame to determine what proportion are currently active members of the population, and to check the accuracy of names, addresses, and other identifying information. While the contractor normally will perform the validity tests, the results should be fully documented for Agency review.

(5) Compare the sample selection procedures in the work plan with the results of the sample selection operations actually carried out at each stage of sampling for the survey proper.

If any sampling is to be done in the field, the contractor should pretest the selection procedures and provide counts of the number of units selected at each stage, broken down by categories for which frame information is available. Agency experts or the contractor should check these counts against the anticipated sample sizes. Frame totals can be checked by (a) applying appropriate sampling weights to the sample counts, and then (b) using tolerances based on estimated sampling errors, comparing them with actual frame totals. Make sure these checks are made before giving the contractor authority to start collecting data for the main survey.

(6) Review the specifications for preparing the sample estimates. Later, when the contractor has completed the preliminary tabulations, check the key statistics against (a) data from prior surveys or other sources and (b) known totals from the sampling frames that were used. (For further details, see "Preparation of the Outputs" in section A of Chapter 6.)

(7) Review the specifications for calculating sampling errors. Check the actual estimates of sampling errors for reasonableness as soon as they are available. An easy way to check estimates of sampling errors for population counts as well as proportions or percents based on these counts is to compare them with the sampling errors that would have been obtained if a simple random sample had been used. The ratios of the contractor's estimates to the corresponding values of the sampling errors for the simple random sample generally range from slightly less than 1 to about 2

or 3, depending on the sample design used. If all the ratios are much larger or smaller, there is likely to be a programming error or an error in the estimation formula (or method).

Another way of checking the reasonableness of the sampling errors is to plot the estimated sampling errors against the corresponding estimates obtained from the sample data (totals, percents, means, etc.). The values usually will follow a fairly regular pattern. Any extreme values may indicate processing errors for the items in question. If the plotted values for a particular class of estimates do follow a regular pattern, a curve can be fit to these calculated values. This curve can be used to estimate sampling errors of items for which sampling errors were not actually calculated.

FOR ADDITIONAL INFORMATION ON SAMPLING --

- Basic Ideas of Scientific Sampling, Second Edition, A. Stuart, Charles Griffin and Co. Ltd., 1976.
- Introduction to Survey Sampling, Quantitative Applications in the Social Sciences, No. 35, G. Kalton, Sage Publications, Beverly Hills, CA, 1983. A few equations, but straightforward and clearly written.
- Sampling in a Nutshell, M. J. Slonim, Simon and Shuster, New York, NY, 1960.
- Survey Sampling: A Non-Mathematical Guide, A. Satin and W. Shastry, Statistics Canada, 1983.

CHAPTER 5

INTERVIEWING

A survey interview is a conversation between an interviewer and a respondent for the purpose of obtaining certain information from the respondent. Coupled with a well-designed, well-tested questionnaire, personal interviews are a powerful, indispensable survey research tool. Whether conducted at the respondent's home or place of business, or over the telephone in a centralized, supervised environment, interviews have been used effectively to collect survey data for more than 30 years. They are especially appropriate for sounding out people's opinions, future intentions, feelings, attitudes, and reasons for behavior, and are adaptable to a wide variety of research situations.

In this chapter we will look at --

- The kinds of quality-assurance procedures the contractor should establish to ensure that their interviewers collect good data from the survey sample;
- The tasks the contractor typically performs to organize and effectively manage the interviewing activities;
- The role of the interviewers in a face-to-face survey; and
- The things the sponsoring office can do to foster the collection of high-quality data.

Our emphasis throughout this chapter is on face-to-face surveys. However, much of the text is relevant to telephone interviewin and, to the extent that interviews are used for follow-up or quality-control purposes, to mail surveys as well.

A. ESTABLISHING THE QUALITY-ASSURANCE PROCEDURES

It is vital for the contractor to establish a set of procedures to assure the quality of the work done throughout the data collection phase. The quality-assurance procedures should cover --

(1) Who specifically is to be interviewed at each sampling unit (or unit of observation). These are called "respondent rules."

(2) How much effort the interviewers should exert to secure an interview. This is established in the so-called "follow-up procedures."

(3) The strategies that are to be used to ensure the collection of high-quality data. These "quality-control strategies" are intended to reduce data errors for which interviewers are primarily responsible and, insofar as possible, to detect and correct these errors.

The respondent rules, follow-up procedures, and quality-control strategies should be incorporated into the work plan and approved by the sponsoring office before any data for the main survey are collected. They should be revised as necessary following any pretests or pilot tests. The contractor should highlight these procedures and strategies in all training programs and instructional materials prepared for the interviewers, supervisors, and support staff.

Let's examine the three types of quality-assurance procedures in greater detail.

1. Respondent Rules

Respondent rules specify the individual or individuals who are eligible, acceptable, or most desirable as respondents for each unit of observation. These rules also specify whether the respondents are to be interviewed alone or with other respondents at the same unit, and whether individuals who are not respondents may be present.

How stringent or flexible the respondent rules should be depends on the kinds of questions to be asked and the conditions under which the interviews are to be conducted (where, when, and the length of the questionnaire). Obviously, the more inflexible the respondent rules, the more "call-backs" the interviewers will have to make to reach the designated respondents. Conversely, the more flexible the rules, the higher the interviewers' completion rates will be.

Respondent rules usually include eligibility criteria such as age (in household surveys) and title or type of responsibility (in business surveys). Sometimes the rules designate only one person in the sampling unit as an acceptable respondent for the unit, e.g., the head of the household, the board chairman, the supervisor of public works. In other cases, anyone who meets the eligibility criteria may be designated as the respondent. For some surveys, the interviewers may be required to talk with several individuals at each unit (all responsible adults, say), with each respondent supplying answers to different parts of the questionnaire. In other surveys, a particular type of respondent may be identified as the "most desirable" respondent, but the interviewer may be allowed to interview any other responsible adult if this person is not available.

Respondent rules also specify whether interviewers may talk with an alternate respondent -- a "proxy" -- after they have made a certain number of unsuccessful attempts to interview the designated respondent(s). Using proxies may produce a marked deterioration in data quality, however. Usually, some information about the units of observation is best supplied by a particular person (the head-of-household, the plant manager). If data are obtained from someone other than the designated respondents, there are likely to be serious gaps, inaccuracies, and biases in the information the interviewer gets. Nevertheless, if it is imperative to obtain some information about the unit of observation, the rules may allow the interviewer to collect data from neighbors, co-workers, or others if the designated respondents cannot be reached.

2. Follow-Up Rules

Follow-up rules prescribe the amount of effort that must be expended to complete an interview with the designated respondent(s) for each sampling unit. Follow-up rules should specify --

- The number of attempts that must be made to secure an interview from a single unit or a cluster of units;
- The time of day the interviewers are to make the initial visit and subsequent visits to each unit; and
- Any allowable deviations from these rules (e.g., to hold down costs, the interviewer may make fewer personal visits to units in sparsely populated areas, if necessary).

For a particular survey, the stringency of the follow-up rules will depend on (a) how vital the researchers believe it is to obtain information directly from the designated respondents rather than proxies; (b) the survey budget (call-backs are costly); (c) how soon the data are needed (inflexible follow-up rules may unnecesarily delay the project); (d) the characteristics of the target population (some types of respondents are difficult to reach during the day); and (e) the characteristics of the areas to be surveyed (e.g., widely dispersed units, inner-city neighborhoods).

3. Quality Control

Guarding against missing and inaccurate data is a major objective in any survey. Strategies must be developed to control three principal types of non-sampling errors that occur during the data collection phase, all of which can seriously compromise the statistics:

- Coverage errors, which result from interviewing ineligible units or failing to interview eligible units;
- Non-response errors, which result when no data or incomplete data are obtained from eligible units (units that should be surveyed); and
- Response errors, which are incorrect reports by the interviewer or the respondent, whether inadvertent or deliberate.

Our concern here is with the effects that interviewing may have on the quality of the data collected in a survey. The fewer errors there are in the data, the higher the data quality will be. Data errors that result from the use of sampling can be measured, and reports of sampling errors included in any reports of the survey results can alert users, so they can take them into account. Non-sampling errors are much more difficult to measure and, therefore, can seriously compromise the survey results.

Non-sampling errors can occur in any survey, regardless of the collection method. Moreover, they do not result solely from poor interviewing. For example, some coverage errors may be directly attributable to the use of incomplete frames, and some non-response and response errors may be the result of poor questionnaire design. In a mail survey where no follow-up interviewing is done, they may be directly attributable to the questionnaire.

However, poor performance by the interviewers or ineffective interaction with respondents can seriously influence the quality of the raw data the interviewers collect, and hence affect the validity of the results. If the interviewers do not adhere to the respondent rules and follow-up procedures, and do not properly administer the questionnaire, the number of non-sampling errors is likely to be very large. Many of these errors may be systematic errors, which no increase in sample size can reduce or eliminate.

Let's examine the sources of (1) coverage errors, (2) non-response errors, and (3) response errors. Then, in the last subsection, look at the principal quality-control strategies survey researchers have developed to reduce these errors during the interviewing.

- Coverage Errors

  The main sources of coverage errors in an interview survey are poorly constructed or outdated sampling frames. For example, the interviewers may be given incorrect listings of the households or businesses they are to cover, so some of the units they attempt to contact are unacceptable, non-existent, or otherwise ineligible. These errors cannot be attributed to the interviewers.

  In some cases, however, the interviewers may be responsible for coverage errors. They may interview the wrong unit by mistake -- because the street number is not clearly marked on the house, for instance. They may even go so far as to make up the answers to a questionnaire for a difficult-to-reach unit, instead of getting data from the designated respondent in that unit.

- Non-Response Errors

  Non-response errors occur, as we said earlier, when the interviewer gets no data (called "total non-response") or incomplete data for an item (called "item non-response") from an eligible sampling unit. Let's look at the sources of these two kinds of non-response errors.

  = Total non-response.

  Total non-response occurs when an interviewer does not obtain any data (or less than the minimum amount required to count as a completed interview) from a sample unit that is eligible for an interview.

Frequently, not all sample units assigned to interviewers are eligible for interviewing. In a household survey, for example, units that turn out to be vacant or demolished are ineligible and will not be treated as non-response cases. On the other hand, where interviews are not obtained for eligible units because of refusals or inability to contact designated respondents, the units will be counted as non-response cases.

It is important that the contract specify in some detail what kinds of units should be defined as ineligible for interview. For example, should households with no English-speaking members be considered ineligible? What about households where all of the eligible respondents are deaf, senile, or otherwise in no condition to be interviewed? These points should be clearly spelled out in the survey contract to avoid later disputes about whether the contracting organization has achieved the target response rate set in the contract. You will recall that we said in Volume I that a response rate lower than 75 percent usually is unacceptable for an Agency-sponsored survey.

Experienced, well-trained interviewers can do a lot to minimize the number of non-responses for eligible units. (See "Locating Respondents" and "Securing Interviews" in section C.) Keep in mind that whatever probability sampling method the contractor uses, every member of the sample must be accounted for if the statistics are to reflect characteristics, opinions, attitudes, etc., of the target population. Therefore, the interviewers must try to complete interviews with all the units or individuals in the sample assigned to them in accordance with the respondent rules and follow-up procedures established for the survey. The level of difficulty they face depends largely on how stringent the respondent rules are, i.e., whether they may interview any responsible adult at the unit or must interview one or more specific individuals.

In addition to total non-response, a partial non-response may occur. Cases are classified as partial non-response if the interviewer fails to obtain acceptable responses to one or more questions but does obtain enough data so the unit need not be counted as a case of total non-response.

The definition of "total non-response" should be included in the contract. This classification normally is assigned to units where responses are missing for any one of certain specified questions or more than a specified number of other items.

= Item non-response.

In what is called "item non-response," the interviewer fails to obtain data for a single item on the questionnaire. Either the respondent or the interviewer may be at fault. For example --

(1) The respondent may remain silent or refuse to answer the question;

(2) The respondent may give an irrelevant answer; or

(3) The interviewer may fail to ask one of the questions or skip to the wrong question, which in either case results in a missing reply.

Interviewers are trained to handle the first two kinds of item non-response with techniques such as pausing briefly to give the respondent time to answer, using words of encouragement to elicit a reply or a more complete reply, repeating questions, probing adequately, and reading questions exactly as they are worded. (See "Asking Questions" in section C for more information.)

- Response Errors

Response errors may be caused by either the respondent or the interviewer. For example --

= Respondents may give inaccurate replies when they do not understand a question and are reluctant to ask the interviewer to repeat or explain it. Or the respondents simply may not know the answer, and rather than appear uninformed or stupid will give a false reply. Or, respondents may deliberately give an inaccurate reply to a question they consider overly sensitive. For example, a 51-year-old man may underreport his age as 47, or overstate his income to impress the interviewer.

= Interviewers may misrecord a respondent's reply (e.g., the same respondent truthfully states his age as 51 but, out of carelessness, the interviewer records it as 41.) Or, interviewers may misread a question, not probe sufficiently when a respondent seems confused or tentative, skip certain questions altogether in the belief they will be able to fill in the answers themselves later when they edit the questionnaire.

Although we said earlier that response errors are caused by the respondent or the interviewer, the underlying cause is the interaction of the two. Other sources contributing to response errors which are not entirely independent of the interviewing process are the conditions of the interview such as the form, content, and wording of the questionnaire; the training and instructions given to the interviewer; and the location of the interview.

The principal things the interviewers can do to minimize response errors are to (a) make an effort to establish a good interaction with the respondent, (b) be faithful to the questionnaire, and (c) maintain an open, neutral position on the questionnaire topics. (See "Asking Questions" and "Recording and Editing the Responses" in section C for details.)

- Quality-Control Strategies

Survey researchers have developed numerous "quality-control" strategies to detect and eliminate or reduce non-sampling errors for which interviewers are primarily responsible. The principal strategies used during the data collection phase to control so-called "interviewer effects" are --

- Monitoring interviewer completion rates
- Observation of interviews
- Preliminary screening of questionnaires
- Validation of interviews
- Reinterviews

QUALITY-
CONTROL
STRATEGIES

Each of these control strategies serves a different purpose. All five should be used in every Agency-sponsored survey where interviewing is the primary collection method, resources permitting. The Agency should require the contractor to specify in the work plan (a) the quality-control strategies that will be used, (b) what each strategy is expected to accomplish, (c) how it will be applied and when, and (d) the procedures that will be used to make sure it is implemented properly.

Let's look briefly at how the five quality-control strategies listed above typically are used to detect and reduce coverage, non-response, and response errors while the interviewing is going on. (Note that in some surveys quality-evaluation strategies may be used at the end of the survey in an attempt to measure the extent of the non-sampling errors. These additional measures are beyond the scope of this Handbook, however.)

= Monitoring interviewer completion rates.

Often a small proportion of the interviewers is responsible for a disproportionate share of the non-response errors in a survey. To help supervisors track the number of these errors each interviewer makes, the interviewers are required to record the specific outcome of each call. For example, to report a (total) non-response for any unit, interviewers must record exactly why they were unable to secure an interview. If a unit is found to be ineligible for interview, the reason must be given.

Interviewers are usually required to prepare a weekly summary of their work, showing the number of assigned cases in four categories: (1) eligible, interview completed; (2) eligible, non-response; (3) ineligible; and (4) pending. Further breakdowns of non-response and ineligible cases, by reason, are often required. Alternatively, these reports may be prepared by supervisors or office clerks, based on the questionnaires turned in by the interviewers.

In either case, these weekly reports should be used by supervisors to monitor the quality and quantity of each interviewer's work. A key indicator of quality is the completion rate -- the percent of all eligible cases for which completed interviews are obtained. Another indicator is the proportion of ineligible cases. A high proportion may indicate that interviewers are misclassifying some eligible units. The average number of call-backs per completed case (those in categories 1, 2, and 3 above) may serve as an indicator of how carefully interviewers are scheduling their calls. Careful review of these and other indicators will allow supervisors to concentrate their attention on interviewers whose work is substandard. (See also the discussion of "Preliminary screening" below.)

- Observation of interviews.

Observation of interviews in both face-to-face and telephone surveys is widely used to train and assess interviewers, and to evaluate respondent reactions in pretest interviews or in exploratory studies.

Direct observation of face-to-face interviews during the survey proper is relatively uncommon, however, because of its high cost. If resources are available for some direct observation of interviewers in the field, supervisors should observe the work of less experienced interviewers and those with below-average performance, as shown by their activity reports and the failure rates of field screenings of their completed questionnaires (see below). A possible substitute for direct observation of face-to-face interviews is to ask each interviewer to tape record one or more of their interviews at specified intervals.

Direct observation of telephone interviews, on the other hand, is relatively inexpensive and therefore a valuable tool for controlling all types of non-response and response errors. It is widely used to monitor and assess telephone interviewers. Throughout the data collection phase, supervisors can easily monitor the interviewer's side of the conversation, quickly correct deficiencies in the way interviewers ask questions, and make sure they ask all the questions. Moreover, with the proper equipment and the permission of the respondent, supervisors can monitor both sides of the conversation and give interviewers valuable feedback on how to improve their skills.

The contractor should develop written evaluation criteria for whatever observation techniques are planned. The criteria are needed to guide the supervisors in which aspects of the interviews they need to look at. Supervisors also should be instructed in how to use the results of their observations to help interviewers improve their performance.

= Preliminary screening of questionnaires.

An initial "field screening" of the questionnaires turned in by the interviewers is an effective way of detecting and correcting many types of non-sampling errors. The term "field screening" is more properly applied to face-to-face surveys, but similar procedures are used by supervisors in conventional telephone surveys to control the quality of the interviews.

Questionnaires may be screened by supervisors or their office assistants. Whoever does the screening should look for (a) missing entries (which may indicate failure to follow skip patterns correctly), (b) inadmissible or questionable entries, (c) unnecessary entries, and (d) illegible entries. The supervisor should record all errors and discuss them with the interviewers.

Field screening may reveal systematic procedural errors by the interviewers, or even faulty instructions or training materials. It is important to detect systematic errors of this type early in the data collection phase, so supervisors can alert the interviewers to their mistakes before too many additional interviews are done. Once the screening has shown that an interviewer is doing good work, it may not be necessary to review all their completed questionnaires -- occasional spot checks may be sufficient.

= Validation of the interviews.

Another important quality-control strategy is for the field staff to verify whether interviewers are actually making all the interviews they claim to have made. Verification is usually accomplished by mailing respondents a card asking (a) if they were interviewed, (b) how long the interview took, (c) if they would be willing to participate again, and (d) if they have any comments or questions about the interview or the interviewer. If a respondent does not return the card within ten days, the supervisor contacts them by phone to verify the interview.

Generally, 10 percent of each interviewer's completed questionnaires are verified each week. Although professional interviewers rarely forge an interview, if any questionnaire fails the verification test the contractor should verify all the interviewer's previous work.

= Reinterviews.

Reinterviews may be an effective method of measuring response errors. They should be done soon after the initial interviews because the longer the interval between the initial review and the reinterview, the more changes in the respondents' characteristics and availability there are likely to be. Sometimes an interviewer with similar training and experience will reinterview the original unit; in other cases, supervisors or more experienced interviewers are used. To minimize the burden on the respondents selected for a "second" interview, usually just a few questions are asked.

The cost of reinterviews is high, however, and the time required to conduct them and process the results -- especially if complete reinterviews are done -- make them unsuitable as a quick, early strategy for measuring interviewer performance. They can be especially useful in continuing surveys, however.

Reinterviews sometimes are used to determine whether units interviewers have called "ineligible" have been correctly classified. Supervisors may reinterview all the housing units in a particular area which interviewers had reported as "vacant." The reinterviews would reveal whether any of these units were actually occupied at the time of the survey. Interviewers sometimes are tempted to misclassify occupied housing units where interviews are inconvenient or difficult to obtain as "vacant," thereby eliminating the requirement to obtain interviews for these units.

B. STAFFING AND ORGANIZING THE FIELD OPERATIONS

In addition to establishing strategies to assure the quality of the data, in a face-to-face or telephone survey the contractor must organize and oversee the work of dozens, perhaps hundreds, of interviewers as well as supervisory and administrative staff.

Although managing the data collection phase of a mail survey is less complex, the contractor must still set up a system to coordinate and control the flow of the questionnaires to and from the respondents. In addition, since mail surveys usually entail some telephone or face-to-face follow-up interviews, staff must be instructed in the proper procedures for these interviews.

In this section we continue our focus on face-to-face interviews, and examine the organizational and administrative tasks a survey contractor typically performs to set up a successful field operation for collecting data in the sampling areas. The four main tasks are --

- Preparing instructions and training materials
- Staffing the field operations
- Training the interviewers
- Coordinating and controlling the field work

ORGANIZING THE INTERVIEWING

Organizing the "field" operations of a telephone survey is similar in many ways, but less complex. There is no need to set up a far-flung field operation as in a face-to-face survey, for example. Usually the interviewers work in one centralized location, supervised by a few members of the contractor's permanent staff. However, instructions and training materials for the supervisors and interviewers must be prepared; the interviewers must be selected and trained; and a system must be set up to coordinate and control the interviewing activities.

The contractor should fully document these procedures in the work plan well before any of the preparatory tasks are initiated. The sponsoring office should review them at the same time as the quality-assurance procedures that we discussed in section A.

1. Preparing Instructions and Training Materials

Once the Agency approves the quality-assurance procedures that will be used to guide the interviewing, the contractor should document them in instructions and training materials for the interviewers, supervisors, and other field staff. How extensive these materials have to be depends largely on the method of collection. Obviously, face-to-face surveys require the greatest number of written materials and mail surveys the least.

Let's look at the three basic guidance documents prepared for a major face-to-face survey: (a) instructions for the supervisors, (b) an interviewer's manual, and (c) a training guide.

- Instructions for the Supervisors

  It is almost impossible to over-emphasize the importance of the field supervisor in controlling the quality of interviewers' work. Yet all too frequently written guidance materials for supervisors concentrate on logistic and administrative matters -- receipt and shipment of materials, payment and allowances for interviewers, etc. These subjects are important, but they do not deal directly with the supervisor's central responsibility, which is to see that the work is done on schedule and that standards of quality are met.

  The instructions to the supervisors should clearly specify --

  = The kinds of quality-related problems requiring communication with the central survey staff, and a well-defined procedure for resolving problems that arise;

  = The quality-control strategies that will be used to assess the work done by the interviewers, and the supervisor's responsibilities in implementing them and evaluating their effectiveness; and

  = A description of the criteria that higher-level field staff or central staff will use to evaluate the supervisor's performance.

- Interviewer's Manual

  A detailed written instruction manual for the interviewers is essential for every survey. Supervisors will also use this manual in their training and for oversight purposes.

If the contractor has developed a standard training manual covering record-keeping, interviewing techniques, and other features common to all surveys, it may be sufficient to prepare a supplement to their standard manual which will cover only the special features of the Agency's survey such as --

- How the respondents were, or are to be selected, and the procedures for locating them;
- The respondent rules;
- The follow-up procedures, especially how to deal with various non-response situations;
- The quality-control strategies to be used;
- The objectives, purpose, and scope of the survey;
- Question-by-question specifications explaining the intent of each question; and
- Any special administrative matters, e.g., the length of the data collection period, who to contact in case of problems, what to do with the completed questionnaires.

Training Guide

A formal training guide for supervisors and others conducting interviewer training sessions is often a desirable supplement to the interviewer's manual. The guide should include topics the trainers should cover, the order in which they are to be taken up, and practice exercises, quizzes, etc., for each training session. The guide can be in outline form or it may be a verbatim guide.

To supplement the training guide, the contractor may develop other materials such as --

- Test exercises, to be completed at various points in the training;
- Written instructions for "mock" interviews;
- Audio-visual materials such as taped demonstration interviews; and
- Slides and other visual aids showing maps of the sampling areas, questionnaire forms, etc.

2. Staffing the Field Operations

Once the instructions and training materials are ready, the contractor must assign existing staff or recruit new staff to carry out the data collection activities. To complete the fieldwork for a major face-to-face survey, normally several dozen interviewers located in 50-100 sampling points (cities or counties), several field supervisors and support personnel, staff for overall project supervision, and a full-time central office will be needed. There should be enough supervisors so they all will have adequate time to monitor the performance of the interviewers assigned to them.

The staff people most directly involved in the fieldwork are (1) the field supervisors and (2) the interviewers themselves. Let's briefly examine their respective responsibilities.

- Supervisors

Some supervision of the interviewers is essential in every survey to detect poor work and assure that the fieldwork proceeds smoothly. Sometimes, centrally-located supervisors direct the work of a mobile field staff, which moves into the various sampling areas. Some survey research firms prefer a network of perhaps a dozen supervisors, who work on a regional basis and move with the field staff from area to area. Whether the field supervisors are centrally located or dispersed, they are the main link between the head office and the interviewers in the field.

The contractor should establish some equitable ratio of interviewers (and other field staff) to supervisors. The ratio should be small enough so the supervisor is able to spend sufficient time both in the field and in the regional (or central administrative) unit to regularly review and evaluate the work of the interviewers for whom they are responsible. The appropriate ratio for any specific survey will depend on factors such as the experience of the interviewing staff, the size of the assignment area, the type of transportation and communication facilities available, and the amount of time the supervisors are required to spend on matters not directly related to the survey.

Each field supervisor is responsible for hiring, training, and maintaining a staff of interviewers in the areas assigned to them. They should be in constant communication with interviewers through personal visits, mail, and telephone contacts.

The field supervisors, along with a support staff of clerical personnel who usually work in the areas where the interviewing is going on, are responsible for --

(1) Arranging travel and lodging for staff and interviewers;

(2) Preparing specific work assignments for the interviewers -- areas, times, lists of households -- or, in the case of a business survey, coordinating and scheduling interview sessions;

(3) Logging in the completed questionnaires and control forms (the interviewers' evaluations, notes, weekly activity reports, etc.);

(4) Scanning the questionnaires for completeness and accuracy, and forwarding them for editing and coding;

(5) Regularly evaluating the interviewers' work, using the quality-control strategies disussed in the previous section; and

(6) Preparing detailed reports on the field activities. These will be used to prepare periodic progress reports for the Agency showing the number of interviews completed or partially completed, the number of refusals, the number of verifications, etc., and the overall response rate.

- Interviewers

In any face-to-face or telephone survey, interviewers play a major role in the quality of the responses and hence in the quality of the results. In some EPA-sponsored surveys, the interviewer is the only link between the contractor's central office staff and the respondents.

No matter what size sample is to be surveyed, the contractor must establish policies and procedures for selecting and training the interviewers and maintaining their morale. A relatively small face-to-face survey of 500 respondents may involve hiring and training as many as 30 interviewers. Keeping interviewer workloads on each survey small will help to (a) keep interviewer travel costs low; (a) minimize the time needed to complete the fieldwork; (c) avoid making the interviewers' job too repetitive and monotonous; and (d) Minimize the effects of systematic errors by individual interviewers.

There is a wide range of practices among survey research firms regarding the hiring of interviewers. Most reputable survey research firms maintain a network of skilled interviewers they can call upon on an as-needed basis. Interviewers usually are recruited on the basic of written applications, followed by a lengthy personal interview and a written test to evaluate the basic clerical skills needed to record, summarize, and edit respondents' answers.

At the end of the project, interviewers generally are rated on their productivity, accuracy, cooperation, and dependability.

Firms typically maintain a file of the names, capabilities, and performance ratings of those who have passed the initial screening. In addition, the file contains detailed information on the interviewers' geographic location, hours available for work, educational background, special skills, current availability, and the results of performance evaluations on previous surveys.

Before hiring interviewers for a specific project, it is important to make sure that they are able to work at the necessary level during specific hours; are able to get to the interview locations; and are willing to work in the assigned areas.

People become interviewers for many reasons. They are motivated by the flexible working hours, the chance to interact with others, and the opportunity to satisfy their curiosity about a variety of research topics.

While there is no such thing as an "ideal" interviewer -- much depends on the nature of the survey, the most sought after qualities typically are intelligence, dedication, honesty, dependability, attention to detail, a professional attitude (neither overly social nor overly aggressive), and an ability to adapt to a variety of interviewing situations (different types of people, different areas, etc.).

Once interviewers are hired, maintaining morale is vital. Good working conditions, a reasonable schedule of assignments, equitable pay rates, and bonuses for high quality work and difficult assignments all contribute to their efficiency.

3. Training the Interviewers

One of the contractor's most important tasks in preparing for a survey is to train the interviewers. The contractor should begin training those who will be used for the main survey shortly after the Office of Management and Budget approves the clearance request.

No matter how skilled or experienced an interviewer or how simple the questionnaire, the interviewers must be --

- Thoroughly instructed in the specific objectives, the rules, and procedures of the survey;
- Taught all quality-assurance procedures they will be responsible for, and the procedures for reporting their progress to the supervisor; and
- Taught a standard format for recording respondent replies.

If the interviewers are inexperienced, they also should be instructed in basic interviewing skills (techniques for gaining entry, probing), and be taught how to plan and update their calling schedules so as to make the best use of their time and travel.

Survey research firms use a variety of techniques to train or retrain interviewers -- interactive lectures, home study programs, practice interviews, and practice in the field. Often a final exam on the field procedures is given as well.

Most face-to-face surveys are complex enough to require interviewers to attend a two-to-five day training conference. These are sometimes held at several different locations around the country. A field supervisor and several professional trainers generally lead the training. Training is guided by the interviewer's manual, the training guide, and various other training aids the contractor has prepared.

The supervisor should evaluate both the effectiveness of the training sessions and, by rating the trainees' performance in practice exercises, quizzes, and exams of various kinds, the extent to which each interviewer has mastered the essentials. Interviewers who are clearly incapable of doing work in the field should be eliminated from consideration, reassigned, or given additional training.

Once the interviewing is in progress, the field staff may provide training for new interviewers or conduct special sessions to reinforce the initial training.

The intent of these combined training techniques is to ensure that the interviewers are capable of collecting complete and accurate data and are fully prepared to elicit respondent cooperation.

4. Coordinating and Controlling the Fieldwork

In addition to hiring and training interviewers, supervisors, and administrative support staff, the contractor must set up a system to coordinate and control the fieldwork. For most surveys, this means establishing procedures for --

- Scheduling and tracking the work of several dozen interviewers for several weeks, or perhaps months.

  Once the contractor has determined how many interviewers will be needed, either the central administrative unit or the field supervisors will prepare a schedule of the units each interviewer must cover. The assignments are based on the interviewer's availability and experience, and often the special characteristics of the sampling areas that have to be covered. For example, although most interviewers are women, if high-crime areas are to be surveyed (particularly at night), male interviewers should be assigned to those areas.

  For both economic and administrative reasons, it is necessary to limit the length of the interviewer's assignments. However, from a practical standpoint, the field supervisors should allow the interviewers enough time to cover all their assigned units and to make whatever number of call-backs were established in the follow-up procedures.

- Controlling the flow of materials to and from the field.

  Once the data collection begins, the pace of the administrative work accelerates rapidly. Unless the contractor establishes close control over the flow of materials to and from the field, chaotic conditions may result. Often a central administrative unit at the contractor's main facility will be given the responsibility of sending instructions and training materials, blank forms and questionnaires and other necessary supplies to field personnel. This same unit also can receive and screen the questionnaires and other such materials completed in the field. A regional field organization frequently is incorporated into the loop. Each unit in the communications chain must maintain accurate records of its own, particularly regarding the response status of each sample unit.

- Resolving problems in the field.

  The contractor must develop a system for the field supervisors to report problems encountered in the field to the regional supervisors or the central administrative unit. If the resolution of these problems affects the existing procedures, all staff should immediately be notified of the changes.

C. CONDUCTING THE INTERVIEWS

Let's turn now from methodological and organizational concerns, for which the researchers, analysts, and administrators on the contractor's staff are responsible, to the practical aspects of interviewing -- the actual conduct of the interviews. We will examine the four principal tasks of the interviewers in a face-to-face survey, which are --

- Locating the respondents
- Gaining respondents' cooperation
- Asking the questions
- Recording and editing the responses

THE INTERVIEWER'S MAIN TASKS

We'll focus our discussion on formal interviews, where the interviewer's goal is to obtain full and accurate answers to a fixed set of items and record them on a standardized survey questionnaire. *When a structured* questionnaire is administered in a uniform way, the researchers and analysts can be reasonably confident that all the answers are comparable. For this reason, formal interviewing is the norm for statistical surveys. This does not mean that formal interviewing allows no flexibility. The interviewer can explain and probe and adjust the speed of the interview -- but within some predetermined limits. Rarely are the interviewers permitted to change the wording or order of the questions, and probing may be allowed only for certain questions.

1. Locating Respondents

In most face-to-face surveys, only about one-third of the interviewers' time is actually spent interviewing. Their most time-consuming pursuit is simply finding the respondents. Approximately 40 percent of an interviewer's time, studies show, is spent traveling and locating respondents. The remainder is devoted to clerical and editing tasks. (Note that in a telephone survey, no time is lost in travel and comparatively little is wasted in searching for the respondents. This is why the cost of a phone survey is about half that of a a face-to-face survey of comparable size.)

How much of the interviewers' time is spent locating the respondents depends largely on the respondent rules.

In a household survey, usually less than half of the interviewer's initial contacts result in completed interviews -- either because no acceptable respondent is home or none of them will agree to be interviewed at the time. Interviewers often have to make several return visits before they secure an interview with an acceptable respondent. If the respondent rules require an interview with one or more specific individuals in the household, a still greater number of call-backs are likely to be necessary. Since the sample units assigned to any one interviewer are often spread over a broad geographic area (a town or county, perhaps), a lot of travel -- and frustration -- are not uncommon.

Locating non-household respondents poses somewhat different problems. Physically locating them usually is not difficult. The main problem in business or industrial surveys is finding the people most qualified to answer the questions. Several call-backs may be necessary before the interviewer locates the right people, and is able to schedule interviews with them.

2. Gaining Respondents' Cooperation

Once the interviewer has located a respondent, the next task is to secure an interview. The way interviewers introduce themselves, the identification they carry, what they say about the survey, how they dress and behave, and the courtesy they show to all the people they come in contact with -- not simply the respondents -- all have a bearing on how successful they are in getting respondents' cooperation. The person the interviewer talks to initially may not be an acceptable respondent, but that person may be able to provide information on when the desired respondent will be at home and ultimately may influence the person's willingness to cooperate.

The interviewer should present a positive, pleasant, relaxed, professional image, and offer the respondent proper credentials -- a picture ID showing the name of the survey research firm they represent, possibly a calling card, and other materials that will demonstrate the integrity of the firm and the importance of the research effort.

The interviewer should briefly explain the nature of the study, the purpose of survey research, and the reasons they want to talk with the respondent. The interviewer also may explain how the data will be used, and who will be permitted access to the data. Explanations about the extent of disclosure of individual responses are especially important to business or industrial respondents, who frequently have strong concerns about revealing trade-sensitive or confidential information.

Most household respondents will agree to be interviewed if approached properly. They do so because they are curious about the subject matter or surveys in general, or because they are pleased to have an opportunity to express their views to someone. Sometimes they agree just because it is harder to say "No" than "Yes" to a skillful interviewer.

Some respondents are willing to be interviewed with only a brief explanation of the purpose of the visit; for others it will be necessary to go into some detail. Respondents have various concerns and questions -- why they were selected, what good will the survey do, why isn't the person next door being interviewed instead -- and the interviewers must give correct and courteous answers.

In no case should an interviewer exert undue pressure to obtain an interview from a reluctant respondent. Responses given reluctantly are likely to be less accurate than those of a more willing respondent. Faced with a persistent refusal, it is best to make no further attempts to get an interview. Sometimes a second approach by the supervisor or a more experienced interviewer will succeed in "converting" a refusal to a completed interview.

Respondents may refuse to be interviewed for any number of reasons -- they are reluctant to break their daily routine; they have other obligations; they are afraid or suspicious of the interviewer; or they are indifferent or hostile to the Federal government, the subject matter, or research in general. Studies show that the respondent's attitude towards surveys in general, based on their own experience and what they have heard from others, is the overriding factor in their decision to grant or refuse an interview.

3. Asking Questions

Once the respondent agrees to be interviewed, the interviewer should immediately try to establish a good interaction so the respondent will be cooperative in supplying the required data. Ideally, the interviewer will have an opportunity to talk with the respondent in private long enough to complete the questionnaire with no disturbances.

As we said at the beginning of this section, the goal of a formal interview is to obtain full and accurate answers to a fixed set of questions. In addition to reading the questions slowly and deliberately so there is no chance they can be misinterpreted, the interviewer should do whatever is necessary to get satisfactory answers. An important part of the interviewer's task, in fact, is to assess the adequacy of the respondent's answers and, if necessary, to take steps to get more information.

Whenever necessary, the interviewer should --

- Ask the respondent if they would like the question clarified or repeated;

- Provide feedback to indicate that an adequate reply has been given or that something else the respondent said has been noted or understood;

- Clarify aspects of the respondent's task which seem to be problematic or confusing; for example, confirm the frame of reference of a particular question;

- Check with the respondent to make sure that a particular response was correctly heard or interpreted;

- Motivate the respondent to complete the questionnaire by interjecting a few words of encouragement from time to time; and

- Control the direction and extent of the respondent's replies, by keeping the respondent from digressing or by reading the next question as soon as a satisfactory answer is recorded, for example.

4. Recording and Editing Responses

Although asking questions well is a critical aspect of a formal interview, the information the respondents provide will be lost if it is not recorded accurately and fully. All interviewers should use the same methods and conventions for recording responses and for editing the questionnaire after the interview is over.

Recording answers may seem to be a relatively simple task, but interviewers sometimes make serious errors. The reason is that interviewing is a fairly tiring, repetitive activity and often a lengthy and complex one as well. In recording replies, interviewers often must follow complex skip instructions and coding rules, and, at the same time, listen carefully to the respondent so they can be ready to take whatever action is necessary to deal with a vague or inadequate reply.

To minimize recording errors, interviewers are trained to check the questionnaire for omissions, ambiguities, illegible entries, and clerical errors before concluding the interview and while the respondent is still available. The interviewer also should note where probes were used, and make a few comments on the interview situation. If a tape recorder is used as a backup in a long interview, the interviewer should transcribe and edit any new information onto the questionnaire.

D. MONITORING THE INTERVIEW PROCESS

As project officer, there are several things you can do, both before and after the fieldwork begins, to foster the collection of high quality data.

Before hiring a contractor, pay particular attention to the following items in the offerors' proposals:

(1) The firm's experience in managing surveys where interviews were used to collect a similar volume of data. Selecting a survey research firm with a good track record in conducting surveys of similar size and scope is usually the best guarantee of getting high-quality data from your survey.

(2) The proposed interviewing activities. Proposals should include clear-cut plans for: (a) quality assurance; (b) selecting, training, and supervising the interviewers and administrative staff; and (c) organizing and overseeing the interviewing activities. We strongly recommend that you have a survey expert review these plans, regardless of what primary collection method the contractor plans to use. Even in a mail survey, some interviewing normally must be done to follow up nonresponse and response errors.

The quality of the data gathered in a face-to-face survey depends largely on the work done by the interviewers. Inaccuracies, omissions, and biases in the data they collect can be kept to a minimum by good training; rigorous use of the quality-assurance procedures established for the data collection; attentive oversight by the contractor throughout the data collection phase; and close monitoring by the sponsoring office.

Therefore, after the contractor is retained --

(3) Have a survey expert review the quality assurance procedures and the procedures for controlling the field operations, as described in the work plan (see sections A and B).

(4) Participate in the pilot test. Go along on some of the interviews as an observer. Attend the interviewer debriefing sessions during and following the pilot test. Work with the contractor on revising the interviewing procedures for the survey proper, if necessary. This will expedite any changes in the questionnaire or the interviewing procedures that require Agency approval. Circulate the pilot test report to survey experts, and make sure the contractor takes proper account of all comments and suggestions before any data are collected in the main survey. (See section A of Chapter 3 for more information on pilot tests.)

(5) Review drafts of all instructions and training materials the contractor prepares for the interviewers and supervisors. Attend as many interview training sessions as possible. You can explain the study goals,

emphasize the Agency's interest in obtaining high quality data, and answer any questions.

(6) Once the data collection begins, make occasional visits to field sites or the facility where the phone interviews are being conducted. If the interviewing is not proceeding according to plan, advise the contracting officer so the Agency can take whatever steps are necessary to correct the problems.

(7) Have a survey expert review the contractor's progress reports during the data collection phase to make sure the contractor is (a) maintaining the schedule, (b) achieving the response rates specified in the work plan, and (c) using the quality-control procedures established in the plan.

FOR FURTHER INFORMATION
ON INTERVIEWING --

- Interviewer's Manual, Survey Research Center, Revised Edition, Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, 1976. Excellent guide to the practical aspects of interviewing.
- Interviewing, Richardson, Dohrenwend and Klein; Basic Books, New York, NY, 1965.
- "Questionnaire Construction and Interview Procedures," Research Methodology in Social Relations, Fourth Edition, A. Kornhauser, P. Sheatsley, and Kidder,et al; Holt, Reinhart, and Winston, New York, NY, 1981.
- Survey Methods in Social Investigation, Second Edition, C. Moser and G. Kalton, Basic Books, New York, NY, 1972. Chapter 12, "Interviewing."
- The Dynamics of Interviewing: Theory, Technique, and Cases, R. L. Kahn and C. F. Cannell, John Wiley & Sons, New York, NY, 1957.

DATA PROCESSING

In most EPA surveys, the contractor is required to process the "raw" data collected from the sample into usable information. Processing involves a series of manual and computerized operations to reduce responses on the questionnaires to machine-readable form so they can be stored, retrieved, summarized, and analyzed.

The desired end-product of these processing operations is a "clean" -- virtually error-free -- data file, usually preserved on magnetic tape. The data file is then programmed by the contractor or the Agency to produce a variety of "output" reports, ranging from simple tables summarizing the characteristics of the data base to highly sophisticated statistical analyses.

In this chapter we discuss --

- The eight fundamental steps in processing survey data; and
- How to monitor the contractor's data processing activities so that the end-product is a clean data file, suitable for preparing tabulations and analyses that will reveal the salient features of the data base.

A. STEPS IN PROCESSING SURVEY DATA

This section examines the eight steps involved in processing the data collected in a typical statistical survey to produce the results for the final report.

- Development of the processing procedures
- Staff selection and training
- Receipt and control of the completed questionnaires
- Manual review and edit
- Coding of open questions
- Data entry
- Error detection and resolution
- Preparation of the outputs

DATA
PROCESSING
PROCEDURES

Preceding page blank

The complexity of these steps in any particular survey depends on three factors:

(1) The extensiveness of the outputs defined in the analysis plan. The analysis plan, which specifies the preliminary tabulations and the types of analyses to be prepared from the data file, not only influences the design of the questionnaire, the sampling plan, and the data collection procedures, but also guides the processing operations. (See Chapter 1 for more information on the analysis plan.)

(2) The size and complexity of the questionnaire. The nature of the questionnaire profoundly influences the processing procedures. If there are many open questions, which require respondents to frame answers in their own words, editing and coding the raw data on the questionnaires will necessarily be more complex. Conversely, if most of the questions offer a fixed range of pre-coded responses, or if a CATI-programmed questionnaire is used, several processing steps may be bypassed.

(3) The size of the sample and the complexity of the sampling procedures. These determine how many questionnaires have to be processed and how much weighting and other treatment of the data are needed to produce results for the final survey report. (See Chapter 4.)

Let's turn now to the tasks the contractor typically will perform during each of the processing steps listed above.

- Step 1: Develop the Processing Procedures

The first step in transforming the raw data that have have been collected from the respondents into usable information is to develop a set of procedures for processing the questionnaire data.

The processing procedures are one of the six components of the work plan. The contractor should develop them after major decisions on the questionnaire, the sampling plan, and the analysis plan have been made.

The data processing procedures should specify --

= The specific tasks the contractor will perform after the completed questionnaires arrive at the central processing facility to produce a clean, virtually error-free data file from which the contractor or the Agency can produce the descriptive tabulations or analytic interpretations of the data base to meet the objectives of the survey;

= The software, hardware, and personnel to be used for each of these tasks;

= Provisions for training processing personnel in the special procedures developed for the survey;

= The quality control techniques that will be used to minimize errors at each step of the processing;

= A flow chart for the tasks to be completed at each step; and

= A complete listing and schedule of the tabulations and other output reports that will be generated in preparation for the analysis.

The sponsoring office may establish some preliminary specifications for the processing operations during the design phase of the survey, particularly the form and content of the tabulations (or desired outputs). Once hired, the contractor will have to work with Agency data processing experts, systems analysts, and subject matter specialists to make sure the computerized output reports are clearly defined. This should be done before any computer programs to generate these reports are written. Normally, existing statistical software packages can be modified to accommodate the Agency's tabulation and analysis requirements. However, if the contractor has to develop any new software, sufficient time and resources must be allowed.

Be sure to have appropriate Agency experts review the final processing procedures before giving the contractor the go-ahead to process any data. If the contractor pretests these procedures -- usually in a pilot test or "dry run" of the main survey -- these experts also should review the adequacy of the preliminary outputs generated from the pilot test data. The contractor should incorporate any modifications they recommend at least two weeks before processing any data collected in the survey proper.

- Step 2: Select and Train Staff

Most of the people who will be involved in the data processing operations will be permanent members of the contractor's staff with experience in processing survey data. For most surveys the staff also will include a data processing manager; a computer center manager; operations personnel; clerical, coding, and editing personnel; an operational control unit; data entry personnel; systems analysts; and programming personnel. Usually a supervisor will be assigned to oversee each step of the processing, e.g., the initial screening of the completed questionnaires, the manual edit and coding, the transfer of the data to machine-readable form, the final computer edit and "treatment" of the data, and the preparation of the tabulations.

No matter how experienced the contractor's professional staff, all processing personnel, especially the editors and coders, should receive formal training in the special procedures developed to screen, edit, and code the survey data. Data entry personnel also need a short training course. The systems analysts and programmers, too, should be thoroughly oriented in the informational and analytical objectives of the survey before their work on the project begins.

For most surveys, the contractor will have to prepare instructional and reference materials to train and guide the editors and coders. These materials typically include procedures for coding each open question and for dealing with omissions, inaccuracies, and inconsistencies in the data (item non-response). They should be updated throughout the data processing phase.

The actual processing of the data (Steps 3 thru 7) begins shortly after the first few batches of completed questionnaires arrive at the processing facility. Appropriate members of the contractor's staff will first check in and screen the questionnaires (Steps 3 and 4) and code any open questions (Step 5). Next, other staff will manually key the data either onto cards or directly into the computer (Step 6). Then comes the final "cleaning" of the data file and the classification and sorting of the data, all of which are operations usually performed by a computer (Step 7). The last task is the preparation of various tabulations and analyses which summarize and interpret the content of the file, along with a report fully documenting the processing procedures (Step 8).

Note that if computer-assisted telephone interviewing is used as the primary collection method, several steps are bypassed because the respondents' answers are keyed directly into an on-line terminal during the interviews. Despite CATI's advantages, it should be used only for large surveys

-- over 300 respondents, say -- because of the high cost of the initial programming.

- Step 3: Screen the Questionnaires

  Since all members of the sample must ultimately be accounted for, strict control of the questionnaires (and other paperwork generated during the data collection phase) is essential. The contractor should assign a control number to each questionnaire. The number is usually placed on the title page. The purpose of the control number is to permit the processing staff to identify data from each questionnaire at any point in the processing.

  During this step, clerks at the main processing facility log in the questionnaires soon after they are returned by the respondents (in a mail survey) or the field supervisors (in a face-to-face or telephone survey).

- Step 4: Review and Edit the Questionnaires

  After logging the control numbers, the clerks will batch the questionnaires and forward them to an editing and coding supervisor for screening. The amount of screening done at this stage of the survey depends on the method of collection and how much screening was done in the "field."

  In face-to-face and conventional telephone surveys, questionnaires often receive a preliminary screening by the field supervisors to rectify obvious problems and errors. However, an additional review by the processing staff is almost always done to check for legibility, completeness, and internal consistency. This is especially critical for the first few batches of questionnaires. The hand screening is an effective way of detecting systematic errors the interviewers or other field staff may be making before the interviewing is too far along. Any questionnaires containing major problems generally are returned to the field supervisor for action.

  Errors on mail questionnaires, on the other hand, are referred to other staff for follow-up action to fill in the missing or inconsistent entries -- usually via phone interviews -- before further processing is done. The purpose of this screening is to isolate questionnaires that --

= Are ready for further processing;

= Contain omissions and inconsistencies requiring some follow-up (usually in short face-to-face or telephone interviews) before further processing is done;

= Will be counted as "non-response" cases because there are too many omissions or illegible answers to warrant follow-up; or

= Are deemed "unacceptable" for processing for other reasons, e.g., the questionnaire was completed for an ineligible unit.

It is essential that you and the contractor fully agree on the precise criteria to be used for the screening operations. Usually, to be considered acceptable for processing, a questionnaire must contain legible and complete responses for all key variables and no more than a specified number of omissions for other items.

The clerks doing the screening also may do a thorough review and edit of the questionnaires or, depending on the complexity of the questionnaires, may forward them to editing or coding specialists.

The purpose of a manual review and edit at this stage of the processing is to catch errors before the data are transferred to punch cards or computer tape. Hand editing is relatively slow and inefficient for catching errors, but in a small survey where the data are relatively complete, it plays a major role in the processing. A subsequent computer edit (also called "machine edit") involving a more detailed and complete application of the editing "rules" is vital (see Step 7). The computer edit also serves to detect and correct human errors introduced during the coding and data entry stages, discussed next.

- Step 5: Code Open Questions

Many EPA survey questionnaires include one or more open questions. These questions may generate a large number of different yet acceptable responses, which must be grouped into a reasonable number of manageable response categories so they can be counted and analyzed. This process is called coding.

Closed questions are usually "pre-coded" (coded beforehand directly on the questionnaire). The codes are often very simple. For example, "Yes" is coded "1" and "No" is coded "2." The codes are printed on the questionnaire in machine-readable form. Fully pre-coded questionnaires thus bypass the manual coding step and the replies are entered directly into the computer.

Codes for open questions often require a lengthy development process. First, the investigators tentatively define a few codes for a set of plausible responses to each open question. The coded response categories then are matched against the answers actually given by respondents in the pretest. Usually the initial codes have to be redefined to fit the pretest responses, and perhaps tested again. After the first 50 to 100 questionnaires in the main survey are edited and coded, the codes may be further refined. Still further adjustments may be made later if the coders have difficulty fitting existing codes to actual responses on new batches of questionnaires that arrive for processing.

The actual coding of the replies to open items may be done by the interviewers (partially-open questions are coded during the interview); their supervisors (shortly after the interviewers turn in the completed questionnaires); or, most frequently, by experienced coders at the processing facility. Whoever does the coding uses a special coding manual listing the codes defined for each open question.

Quality control of the coders is vital. The work of each coder must be checked periodically for accuracy and consistency with the codes defined in the manual. Processing supervisors normally check 100 percent of each coder's work at the start. Because coding errors tend to decrease as the clerks become more familiar with the subject matter, a random sample -- usually 10 percent of the coded questionnaires -- is checked after the coders' errors decline to an acceptable level.

To control consistency among the coders, supervisors periodically run tests on a sample of the coded questionnaires and establish a "rate of agreement" for each question. Typically the rate is based on the number of times pairs of experienced coders select the same code for a particular response.

• Step 6: Enter Data

The next step in the processing is to transfer the edited and coded data from the questionnaires onto a computer tape, a disk, or some other machine-readable medium.

The two most common methods of transferring (entering) data are (1) to keypunch them onto cards or (2) to key them directly onto tape or disk through on-line terminals connected to a computer. Both methods involve manual keying and are, therefore, subject to human error.

When the keypunch method is used, two different operators keypunch one or more cards containing the data from a single questionnaire. Quality control is achieved by a computer-assisted comparison of the cards to spot and reconcile any differences.

Direct keying (key-to-tape or key-to-disk) is rapidly replacing keypunching as the preferred method of data entry because direct keying is more efficient and more convenient. In direct keying, experienced operators type data from the questionnaires at entry stations that have a keyboard similar to a typewriter. Quality control is achieved through periodic checks of the operators' output as well as the data entry equipment and software. Some of the newer key-to-disk equipment can be programmed to identify (and in some cases correct) inadmissible values or codes.

There is another still more sophisticated method of data entry called optical scanning. A scanner "reads" the data on the questionnaire and enters them directly to the computer medium. Optical scanning still is not widely used for processing survey data, but it undoubtedly will enjoy broad application in the future.

• Step 7: Detect and Resolve Errors in the Data File

The next step is to "clean" the data to enhance its quality and facilitate the subsequent production of tabulations and analyses. Data cleaning is the process of detecting and resolving inaccuracies and omissions in the data file. Often it is the most complicated and time-consuming step of the processing.

In almost all surveys today, the bulk of the work of detecting and resolving data errors is performed by a computer. First an intensive machine edit is performed to identify inaccuracies and omissions, and then various techniques are used to correct or convert unacceptable entries into a form suitable for tabulation and analysis.

= Computer edit.

In a computer edit, the first step is to program the computer to check for inconsistent or "impossible" entries, some of which may have been introduced in the previous processing steps. For example, the computer may be programmed to identify errors such as --

(1) Inadmissible codes -- the code attributed to an item does not correspond with the permissible replies in the coding manual (a code "4" has been entered for an item to which only codes "1" and "2" have been assigned);

(2) Out-of-range entries -- the amount that has been entered is below or above the permissible values programmed for that item;

(3) Omissions -- no entry has been made;

(4) Inconsistencies -- entries for two or more items are not consistent with each other (a respondent is reported to be 14 years old and a physician);

(5) Math errors -- the total for a list of items should be equal to the sum of the amounts shown for individual items on the list.

The computer may be further programmed to print an error message indicating the nature of the failure, or even to correct certain errors and log them.

Decisions on how much editing should be done by hand and how much by machine depend on many factors. For some surveys, several manual checks as well as computer runs using special check-and-edit programs may be necessary to achieve an acceptable error rate. Generally speaking, the more complex the questionnaire, the more difficult it is to develop computer programs for detailed edit-checks; thus considerable manual editing may have to be done. Larger sample sizes tend to make computer editing a more cost-effective option.

= Error resolution.

The computer edit detects errors but does not resolve them. Several techniques are used to deal with the errors the computer has identified. Survey researchers use several techniques to deal with data omissions and inaccuracies in individual questionnaire items (so-called "item non-response"). The principal ones are (1) returning to the original questionnaires to see if errors were made in entering the data or if it is possible to infer correct responses from other information on the questionnaires, (2) having the computer impute values for missing responses, and (3) creating separate categories to report all missing replies. More specifically --

(1) Consulting Questionnaires

Generally, the most reliable procedure for resolving omissions and inconsistencies in the data file is to consult the questionnaires. Data entry clerks sometimes pick up data from the questionnaires incorrectly. Or, if the respondent has left an answer-space blank, it sometimes is possible to infer the correct answer from other information on the questionnaire. Footnotes or write-in comments also may provide helpful information.

For instance, if respondents fail to state their ages, researchers may be able to infer their correct ages from other information on the questionnaires such as dates of birth or school attendance. Inconsistent responses sometimes can be resolved by considering the whole range of information supplied by a respondent and deciding which of the conflicting entries is most plausible, e.g., from information on the income, education, and marital status of the "14-year-old physician" in the example on the previous page, it might be reasonable to assume that the respondent is really 41 years old.

Consulting questionnaires as a means of resolving errors, however, is time-consuming and not always productive.

(2) Imputating Missing Values

Another method of error-resolution is to try and compensate for the non-response bias by having the computer impute values for the omitted and inconsistent replies. Imputation involves assigning values for missing or unusable responses by drawing on information from other sources such as answers to other items on the same questionnaire, another questionnaire from the same survey, or external sources (administrative records or another survey). Imputation corresponds to the weighting adjustments for total nonresponse, which we will discuss in Step 8.

Imputation generally is a faster and less costly error-resolution technique than consulting questionnaires, but it must be used with discretion. Imputed items should be flagged in the data file so that tabulations and analyses can be prepared with and without the imputations, if desired. Also, any reports about the survey should indicate the extent of the imputation so that anyone using the data later can distinguish between real and imputed values.

The extent to which the contractor intends to impute values for missing or omitted replies should be specified in the data processing procedures submitted with the work plan.

Note that the contractor should aim to get good data from the respondents in the first place, and make data adjustments judiciously and strictly as a back-up measure. Imputation can be kept to a minimum by instructing interviewers to carefully check the questionnaires immediately after each interview; regular, thorough, and timely checks of the interviewers' work during the data collection phase; and follow-ups of respondents in mail surveys.

(3) Creating Categories for Unreported Responses

If attempts to resolve omissions and inconsistencies in the data file using the above two techniques are unsuccessful, the researchers may allow the errors to stand and report them as such in the tabulations. For example, they may report a total for all respondents who provided no valid income data in a new category called "income unknown."

Decisions on whether to impute values for omitted and inconsistent replies or to add "not reported" categories in the tabulations depend on a number of circumstances. Using a "not reported" category for tabulating data on such basic characteristics of the sample as "sex" and "age" creates serious problems in the analysis. Analysts sometimes handle this by imputing values for fundamental demographic variables for which considerable related information is available, and creating "not reported" categories for describing and relating information on all others.

- Step 8: Prepare the Outputs

The final step in processing survey data is to prepare the tabulations and other outputs called for in the work plan. The contractor's main tasks at this step are to (1) weight the sampled elements to produce the estimates (the results), (2) prepare the preliminary tabulations describing the data base (the content of the data file) and finalize the analysis plan, (3) apply the procedures described in the sampling plan for calculating the sampling errors, and (4) document the procedures used in preparing the data file.

Let's examine these tasks more closely.

(1) Weighting the Sampled Elements

The first task in generating the tabulations is to weight the virtually error-free data file prepared in the previous step. Except for simple lists of data items, these preliminary reports summarizing the content of the file should be based on weighted data. Weights (or multipliers) are assigned to survey data for three reasons --

= To account for the probabilities used in selecting the sample from the target population.

If all units in the sample have the same probability of being chosen, the survey analysts can obtain valid estimates of some statistics such as proportions, percents, means, and medians without weighting the data. However, to estimate totals, all units must be weighted by the reciprocal of the sampling fraction. For example, if the sampling fraction was 1 in 200, all sample values or totals must be multiplied by 200. If the selection probabilities were not the same for all the units, appropriate weights must be applied to estimate any statistic. (See section D of Chapter 4 for more information on weighting.)

- To adjust for sampled units for which no data were obtained (non-response).

  There are two methods of making adjustments for non-response:

  One way is to increase the weights applied to individual units that did respond and are similar (based on data available for all the sample units) to those for which no data were obtained. For example, if one sample household in a block did not respond, one of the households for which data were obtained would be selected at random and given an additional weight of "2."

  The other way is to apply a uniform weight to all the units in the sample or to those in a particular subgroup. For example, in a business survey, if 20 percent of the sample establishments with fewer than 10 employees did not respond, a weight of 1.25 (100 divided by 80) would be applied to all establishments that did respond.

- To apply more sophisticated estimation procedures such as ratio or regression estimates.

  These procedures require a determination of relationships between variables or the introduction of independent data from other sources, e.g., current population estimates.

The overall weights the analysts ultimately assign to the data will reflect the combined effects of these three types of adjustments. Deciding on the sequence and procedures for weighting the data in a particular survey requires a good technical grasp of the sample design and the data processing system. Sampling and data processing experts at the Agency and on the contractor's staff should work out the weighting and estimation procedures long before the processing starts. These procedures should be critically reviewed by systems analysts at the Agency before the contractor processes any data collected in the survey proper.

(2) Preparing the Preliminary Tabulations

After the weighting and estimation procedures are completed, a data file suitable for generating the preliminary tabulations should result.

Using a standard computer software package or software specially designed for the survey, the contractor then can program the data file to generate a set of preliminary tabulations, which normally will include --

- Frequency distributions(sometimes called "marginal tabulations") of responses for categorical variables (those based on questions with fixed response categories);
- Some simple cross tabulations;
- Estimated totals, ranges, and means (or medians) for the entire target population and for various subgroups;
- Listings of individual responses for selected items, especially for large sample units; and
- Tabulations of key variables showing the number of units for which an item was imputed and how much of the total was imputed, where applicable.

The preliminary tabulations will give you and the contractor an opportunity to review the data base in an organized fashion, and thereby get an idea of its structure and quality before the contractor prepares the final tabulations.

Subject-matter specialists should carefully study these preliminary tabulations before the contractor prepares a revised list of the final tabulations to include in the analysis plan. The list should include the computerized output reports (tables and graphs) that will be prepared to fully describe the content of the data base.

There is no clear line between the output reports generated at the conclusion of the processing phase and those developed for the analysis. However, the analysis of the data base usually goes beyond simple descriptive summaries and explores the underlying relationships among the study variables.

A host of sophisticated analytic techniques may be used to reveal the full informational content of the data base.

Usually, the final tabulations include --

- Detailed descriptive statistics (frequency distributions and cross-tabulations);
- "Measures of central tendency" (means, medians, and modes);
- "Measures of variability" (standard deviations, ranges); and
- Other analytical statistics such as correlations and regression coefficients.

The revised analysis plan should specify for each tabulation (a) the data sources to be used, (b) the variables to be cross-classified, (c) the subpopulations to be included, (d) the statistics to be shown, (e) how the data are to be weighted, (f) the title, subheadings, and footnotes; and (g) the layout. The analysis plan should also include --

- A full description of the methods for quantifying all relevant variables;
- Values of sample weights and all necessary formulas for estimating population means, medians, and variances;
- A list of hypotheses and the tests to be used to evaluate them;
- Descriptions of the variables and respondent groups that may be interrelated, and recommendations for regression and discrimination analyses based on the relationships; and
- Suggested methods for handling problems during the subsequent analysis, that arise from missing data or non-response problems.

You should work with data processing and systems analysts both at the Agency and on the contractor's staff in defining these specifications for the final analysis plan.

(3) Finalizing the Computations of Sampling Errors

The actual calculation of sampling errors for various estimates should be an integral part of the processing operations. Making these calculations after the preliminary tabulations are generated is generally much more difficult, time-consuming, and costly.

The estimates of sampling errors (variances) serve two purposes --

- They may help evaluate the data base. For example, unusually large sampling errors for some items may indicate processing errors; and
- They are essential for determining whether observed relationships are statistically significant or may be due to random variation introduced by the use of sampling.

As discussed in Chapter 4, sampling errors usually are not calculated for *all* the statistics produced from the survey. This is usually unnecessary and often too costly. The contractor's analysts and sampling specialists should select the items for which sampling error estimates are needed, making sure to include all key statistics and a representative set of other types of statistics that are to be tabulated from the data file. (For more details on calculating sampling errors, see section D of Chapter 4.)

(4) Documenting the Processing Operations

Once the final tabulations are completed, the contractor should create a file documentation manual describing the procedures used to edit, code, and weight the data. The manual should identify the source of each data item (on the questionnaire or other document used during the data collection phase) and its position on the file.

If EPA is to analyze the content of the data file, the contractor should submit the documentation manual, the final analysis plan, and whatever other materials (computer cards, for example) Agency analysts will need to study and interpret the data file.

On the other hand, *if the contractor is to do the analysis*, the documentation manual should be submitted for EPA review and approval along with the final analysis plan before the data are analyzed.

A discussion of data analysis is beyond the scope of this Handbook. To assist you in this regard, we have provided a list of excellent sources at the end of this chaper, along with a number of selections offering additional guidance on data processing issues.

The final step of the survey, the presentation of the results, any necessary background information, and the conclusions drawn from the results, is covered in Chapter 8 of Volume I. Often the survey contractor is required to prepare both a non-technical report for the public and a detailed account of the technical findings. Remember that any report about the survey should be issued by the Agency, not the contractor.

B. MONITORING THE PROCESSING ACTIVITIES

Throughout this Handbook we have emphasized that EPA's major impact on the successful outcome of a contract survey comes long before the data collection and data processing activities are under way. Achieving a clean data file on which to base the analytic work is largely dependent on the professional, clerical, and management capabilities of the firm the Agency hires to conduct the survey. As in the data collection phase, the sponsoring office has only limited control over the data processing activities.

Therefore, before the contractor is hired, you should --

(1) Require the offerors to specify in their proposals --

- The formal quality-control procedures they intend to use at each step of the processing;
- How they intend to keep coding and other errors to a minimum; and
- How they will report production and error rates for each step of the processing.

(2) Specify the format and any special requirements for the completed data file to ensure compatibility with other EPA data files and otherwise facilitate the analysis.

(3) Require Agency approval of the key deliverables of the data processing phase (the data file, the tabulations, the estimated sampling errors, and the documentation of the processing procedures). If the Agency is to do the analysis, specify that EPA must approve these deliverables before the contract is closed out. If the contractor is to do the analysis, do not let the contractor begin it until you have reviewed and approved the above products of the data processing phase.

Other things you can do after the contractor is aboard to help assure the quality of the data file and the other deliverables are --

(4) Make sure the questionnaire is designed to facilitate the processing operations.

(5) Before data for the main survey are collected, carefully review the processing procedures and tabulations specified in the work plan. If necessary, work with the contractor on specifications for the content and format of the final tabulations. If a pilot test is done, review the procedures and tabulations and make sure the contractor makes any necessary modifications before processing any data from the survey proper.

(6) Participate in the development of response codes and procedures for treating non-response and "unacceptable" responses.

(7) Scrutinize all progress reports submitted during the processing to make sure the contractor is (a) adhering to the schedule and budget and (b) following the verification and quality-control procedures specified in the work plan.

(8) Have Agency statisticians, project personnel, and data processing experts review the preliminary tabulations and the file documentation manual. All tables should be reviewed to be sure that (a) they are internally consistent; (b) the estimates appearing in more than one table agree; (c) significant changes from comparable data in earlier surveys are adequately explained; and (d) the estimates are "reasonable" based on expectations and data from other sources.

(9) Finally, if the Agency is to do the analytic work, make sure that all deliverables are in good order before the contract is closed out.

FOR MORE INFORMATION ON DATA PROCESSING --

- *National Household Survey Capability Programme, Survey Data Processing: A Review of Issues and Procedures*, United Nations, Department of Technical Cooperation for Development and Statistical Office, New York, NY, 1982.
- *Survey Methods in Social Investigation*, Second Edition, C. A. Moser and G. Kalton, Basic Books, Inc., New York, NY, 1972. Chapter 16, "Processing of the Data," and Chapter 17, "Analysis, Interpretation and Presentation."
- *Survey Research Practices*, G. Hoinville, R. Jowell, and associates, Heinemann Educational Books, London, England, 1978. Chapter 8, "Data Preparation."
- *The Sample Survey: Theory and Practice*, D. P. Warwick and C. A. Lininger, McGraw-Hill, New York, NY, 1975. Chapter 9, "Editing and Coding," and Chapter 10, "Preparation for Analysis."

FOR MORE INFORMATION ON STATISTICAL ANALYSIS --

- *A Guide for Selecting Statistical Techniques for Analyzing Social Science Data*, Second Edition, F. M. Andrews, et al, Institute for Social Research, University of Michigan, Ann Arbor, MI, 1981.
- *Applied Regression Analysis*, Second Edition, N. Draper and H. Smith, John Wiley & Sons, New York, NY, 1983.
- *Searching for Structure*, Revised Edition, J. A. Songuist, E. L. Baker, and J. N. Morgan, Institute for Social Research, University of Michigan, Ann Arbor, MI, 1974.
- "Standards for Discussion and Presentation of Errors in Survey Census Data," *Journal of the American Statistical Association*, Vol. 70, No. 351, Part II, M. Gonzalez et al, September 1975.
- *Understanding Robust and Exploratory Analysis*, D. Hoaglin et al, John Wiley & Sons, New York, NY, 1983.

GLOSSARY

BIAS - The difference between the survey estimate, averaged over repeated samples, and the true value. Sampling bias can result from use of a non-probability sample or from errors in the execution of a probability sample design. Non-sampling bias can result from many factors such as use of an incomplete sampling frame (coverage bias), non-response in the survey (see NON-RESPONSE BIAS), a poorly designed questionnaire, respondent errors, interviewer errors, or processing errors.

BURDEN - In the 1980 Paperwork Reduction Act, "burden" is defined as the amount of time required to collect data from the public using a particular data collection instrument (a questionnaire). The response burden of a particular survey questionnaire is the estimated number of hours each respondent needs to complete the instrument, multiplied by the total number of people to be surveyed. The total number of burden hours for a survey questionnaire must be reported to the U.S. Office of Management and Budget (OMB) if data are to be collected from more than nine members of the public. OMB is responsible for overseeing Agency compliance with the PRA.

CATI (computer-assisted telephone interviewing) - A relatively new method of telephone interviewing in which a structured questionnaire is programmed into a computer, rather than printed on a form. The interviewer sits before a video terminal and asks the questions as they appear on the screen. The interviewer then enters the respondent's replies directly into the computer via a keyboard attached to the terminal.

CLOSED QUESTIONS - Questions offering respondents two or more alternative answers, either explicitly or implicitly, e.g., Yes/No, Male/Female, Strongly Agree/Agree/Disagree/Strongly disagree. When more than two choices are offered, closed questions are sometimes called "multiple choice questions."

CODING - The processing of survey answers into numerical form for entry into a computer, so that statistical analysis can be performed. Coding of alternative responses to closed questions (see CLOSED QUESTIONS) can be performed in advance so that no additional coding is required. This is called "precoding." If some items are precoded or keyed directly (numerical amounts), then coding refers only to the coding of open questions (see FIELD CODING).

Preceding page blank

DEBRIEFING - A meeting of interviewers, supervisors, research analysts, etc., immediately after a pretest or during the early stages of the data collection phase of the main survey. Debriefings alert project personnel to problems with the questionnaire, so they can be corrected before the rest of the interviews are done.

DEMOGRAPHIC CHARACTERISTICS - The basic variables used by survey researchers to classify population groups, e.g., sex, age, marital status, race, ethnic origin, education, income, occupation, religion, and residence.

DEPENDENT/INDEPENDENT/INTERDEPENDENT VARIABLES - Dependent variables are the behaviors or attitudes whose variance the researchers are attempting to explain. Independent variables are those variables used to explain the variance in the dependent variables. Variables such as "occupation" or "income" may be dependent or independent, depending on the purposes of the research and the model used. In more complex models, variables may be interdependent; that is, variable A is affecting variable B while, simultaneously, variable B affects variable A.

DIARIES - Written records kept by respondents to keep track of events that may be difficult to recall accurately later. Diary-keepers are requested to make entries immediately after an event occurs. Sometimes they are compensated with money or gifts for their efforts.

FACE-TO-FACE INTERVIEWS - One of the three traditional interviewing methods used to collect statistical data. In face-to-face interviewing, a trained interviewer poses questions in the presence of the respondent.

FIELD CODING - The coding of responses to open questions by the interviewer during the interview. When this technique is used, the questionnaire includes a set of preprinted, coded replies. Instead of writing down the respondent's answer verbatim, the interviewer checks the preprinted reply that most nearly matches the respondent's reply.

FIELD TEST - See PRETEST and PILOT TEST.

FOCUS GROUPS - An exploratory interviewing technique involving small, informal group discussions "focused" on selected topics of concern to the researchers. The discussions are led by a moderator knowledgeable about the subject matter. The participants are selected from the target population or a specific subgroup of the target population.

FRAME - The source or sources from which the survey sample is drawn. The sampling frame may consist of one or more lists of individuals or organizations, but it also may be a set of city blocks, a set of telephone exchanges, etc.

IMPUTATION - The process of replacing missing or unusable information with usable data from other sources such as responses to other items on the same questionnaire, another questionnaire from the same survey, or external sources (another survey or administrative record). The use of imputation techniques is rapidly expanding in scope and sophistication due to advances in computer technology.

INTERVIEWER INSTRUCTIONS/DIRECTIONS - Instructions to interviewers regarding which questions to ask or skip, how to enter responses, and when to probe (see PROBES). Interviewer instructions are printed on the questionnaire but not read to respondents.

LOADED QUESTION - A question worded in a way that increases the likelihood of a particular kind of response. Loaded questions may legitimately be used to overcome respondent reluctance to report sensitive information. Poorly written questions using "loaded" words or expressions may inadvertently produce biased responses.

MULTIPLE-CHOICE QUESTIONS - See CLOSED QUESTIONS.

NON-RESPONSE BIAS - Non-response bias results when units who do not respond to the survey differ significantly from those who do respond. It can also result from non-response to individual items on the questionnaire.

OPEN (OR OPEN-ENDED) QUESTIONS - Questions allowing respondents to answer in their own words. The open format encourages respondents to express themselves in language that is comfortable to them. Some open questions are coded during the interview using a fixed set of response categories (see FIELD CODING).

PILOT TEST - A small field test replicating the field procedures proposed for the main survey. Usually a purposive sample of 10 to 50 members of the target population is used for the test. A pilot test is more elaborate than a pretest (see PRETEST) in that the proposed collection procedures as well as the questionnaire are tested. Its purpose is to alert the researchers to any operational difficulties not anticipated during the planning and pretesting stage. (Note that some researchers use "pretest" and "pilot test" synonymously.)

PRECODING - See CODING.

PRETEST - A small field test of the questionnaire proposed for the main survey. Usually a purposive sample drawn from various subgroups of the target population is used. Pretests are vital for all Agency-sponsored surveys involving new topics or populations. (Also, see PILOT TEST.)

PROBABILITY SAMPLE - A sample drawn in such a way that each unit (person, household, organization, etc.) in the target population (see TARGET POPULATION) has a known, non-zero probability of being included in the sample. This method of selecting the survey respondents makes possible statistically valid inferences about the entire population the sample is designed to represent.

PROBES - Questions or statements used by the interviewer to obtain additional information from the respondent when the initial answer appears incomplete. Examples of probes are: "How do you mean?" "In what way?" or "Could you explain that a little?"

QUESTIONNAIRE - The complete data collection instrument used by an interviewer or respondent during a survey. The questionnaire includes not only the questions and spaces for the answers, but also interviewer or respondent instructions and an introduction. The questionnaire usually is printed, but recently nonpaper versions are being used on computer terminals (see CATI).

RANDOM DIGIT DIALING (RDD) - A method used to select samples for telephone surveys by random selection of telephone numbers within working exchanges. This method permits coverage of both listed and unlisted telephone numbers.

RANDOM SAMPLE/NON-RANDOM SAMPLE - In practice, the term "random sample" is often used loosely to mean any kind of probability sample. "Simple random sample" is a technical term for a sample in which each unit in the population has the same probability of selection and in which all possible samples of a given size are equally likely to be selected. The term "non-random sample" is used to mean any sort of non-probability sample such as a quota sample, a convenience sample, or a judgment sample.

RECORDS - Documents used to reduce memory error on factual questions. Memory errors are unintentional errors in respondent reports caused by forgetting or incorrectly recalling events or details of events. Examples of records are bills, checkbook records, cancelled checks, and inventory accounts.

RESPONSE BURDEN - See BURDEN.

RESPONSE EFFECTS - Variations in the quality of data resulting from the process used to transmit information from the respondent to the interviewer (where applicable) and ultimately to the data user. The principal sources of variation in quality are the interviewer's performance, the respondent's performance, and the nature of the data requirements and collection methods established by the survey designers.

SAMPLING - Selection of some of the units (a sample) from a population (see TARGET POPULATION) to otain information that that can be used to characterize or describe the whole population. Probability sampling is the prescribed method for Agency surveys. See PROBABILITY SAMPLE.

SCALE QUESTION - A multiple-choice question that asks respondents to rate a particular quality in themselves or some other person or thing. For example, they may be asked whether they agree or disagree with a statement of opinion, about the frequency of a type of behavior, or whether they like or dislike a certain product. Some scales are entirely verbal (sometimes referred to as "fully-anchored scales"), e.g., "excellent," "very good," "fair," "poor."

SELF-ADMINISTERED QUESTIONNAIRE - A questionnaire requiring respondents to read and answer the questions themselves. Self-administered mail questionnaires are one of the three traditional methods of collecting survey data. Note that a questionnaire can be considered to be self-administered even if an interviewer is present to hand it out, collect it, and clarify questions.

SKIP INSTRUCTIONS - Directions on the questionnaire to show the person completing the form which question to ask or answer next, based on the answer to the previous question. Skip instructions make it possible to use a single questionnaire for many different types of respondents because they need answer only those items that are relevant.

SOCIAL DESIRABILITY/SOCIAL UNDESIRABILITY - This refers to the perception by respondents that the answer to a question will enhance or hurt their self-image in the eyes of the interviewer. Examples of socially-desirable behavior are voting, being well informed, and fulfilling moral and social responsibilities. Examples of socially undesirable behavior include alcohol and drug abuse, deviant sexual practices, and traffic violations.

STATISTIC - A summary measure derived from sample data. "Statistics" (plural), in everyday language, refers to a collection of numerical data. "Statistics" (singular) is an academic discipline concerned with methods of converting numerical data into information useful for scientific research, business decision-making, and other similar purposes.

STRUCTURED/UNSTRUCTURED QUESTIONNAIRES - Structured questionnaires specify the wording of the questions or items and the order in which they are asked. They are used for all statistical surveys, regardless of whether the questionnaire is administered by interviewers (in person or by telephone) or by the respondents themselves. Unstructured questionnaires are essentially topic outlines in which the wording and order of the questions are left to the interviewer's discretion. Unstructured survey questionnaires are used primarily in exploratory research for in-depth individual interviews or focus group studies.

SENSITIVE QUESTIONS - These are questions that are likely to make respondents feel uneasy or threatened and to which they may be reluctant to respond. They include questions about socially desirable and socially undesirable activities (see SOCIAL DESIRABILITY/SOCIAL UNDESIRABILITY). For businesses, sensitive questions include those covering information which they may not want to reveal to their competitors or to government regulatory authorities.

TARGET POPULATION - The complete set of people, households, organizations, businesses, or other units that is of interest and from which the samples for pretests and the main survey are drawn.

TELEPHONE INTERVIEWS - One of the three major methods of collecting statistical data. Data are obtained using a structured telephone interview. As in face-to-face interviewing, the interviewer both asks the questions and records the responses. A relatively recent innovation in telephone interviewing is computer-assisted telephone interviewing. (See CATI.)

VALIDATION - The process of recontacting respondents to determine whether an interview was actually conducted. In a broader sense, "validation" also refers to the process of obtaining data from other sources to measure the accuracy of respondent reports. Validation may be at either the individual or group level. Examples include the use of financial or medical records to check on reports of assets or health care expenditures. Unless public records are used, validation of individual responses usually requires the consent of both the respondent and the custodian of the records.

VARIABILITY/VARIANCE - Used in reference to a population, variability refers to differences between individuals or groups in the population, usually measured as a statistical variance or simply by observing the distribution of values for the group. In samples, variability has the same meaning with respect to members of the sample. For estimates based on samples, variance refers to differences between estimates from repeated samples selected from the same population using the same selection procedures. For statistical definitions of variance, see any statistics textbook.

VARIABLES - See DEPENDENT/INDEPENDENT/INTERDEPENDENT VARIABLES.

LIST OF RECOMMENDED SOURCES

- A Guide for Selecting Statistical Techniques for Analyzing Social Science Data, Second Edition, F. M. Andrews, et al, Institute for Social Research, University of Michigan, Ann Arbor, MI, 1981.
- Applied Regression Analysis, Second Edition, N. Draper and H. Smith, John Wiley & Sons, New York, NY, 1983.
- Approaches to Developing Questionnaires, Statistical Policy Working Paper 10, Statistical Policy Office, Office of information and Regulatory Affairs, OMB, Washington, DC., 1983.
- Asking Questions: A Practical Guide to Questionnaire Design, S. Sudman and N. Bradburn, Jossey-Bass, San Francisco, CA, 1982.
- Basic Background Items for U.S. Household Surveys, R. Van Dusen and N. Zill, Social Science Research Council, Washington, DC., 1975.
- Basic Ideas of Scientific Sampling, Second Edition, A. Stuart, Charles Griffin and Co. Ltd., 1976.
- General Social Surveys, 1972 - 1982: Cumulative Codebook, National Opinion Research Center, University of Chicago, Chicago, IL, 1952.
- Interviewer's Manual, Revised Edition, Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, 1976.
- Interviewing, Richardson, Dohrenwend and Klein; Basic Books, New York, NY, 1965.
- Introduction to Survey Sampling, Quantitative Applications in the Social Sciences, No. 35, G. Kalton, Sage Publications, Beverly Hills, CA, 1983.
- Mail and Telephone Surveys: The Total Design Method, D. A. Dillman, John Wiley & Sons, New York, NY, 1978.
- Measures of Social Psychological Attitudes, Revised Edition, J. Robinson and P. Shaver, Institute for Social Research, University of Michigan, Ann Arbor, MI, 1973.

Preceding page blank

- *National Household Survey Capability Programme, Survey Data Processing: A Review of Issues and Procedures*, United Nations Department of Technical Cooperation for Development and Statistical Office, New York, NY, 1982.
- "Questionnaire Construction and Interview Procedures," *Research Methodology in Social Relations*, Fourth Edition, A. Kornhauser, P. Sheatsley, and Kidder, et al; Holt, Rinehart and Winston, New York, NY, 1981.
- *Questionnaire Design and Attitude Measurement*, A. Oppenheim, Basic Books, New York, NY, 1966.
- *Sampling in a Nutshell*, Morris J. Slonim, Simon and Shuster, New York, NY, 1960.
- *Searching for Structure*, Revised Edition, J. A. Songuist, E. L. Baker, and J. N. Morgan, Institute for Social Research, University of Michigan, Ann Arbor, MI, 1974.
- "Standards for Discussion and Presentation of Errors in Survey Census Data," *Journal of The American Statistical Association*, Vol. 70, No. 351, Part II, M. ,onzalez et al, September 1975.
- *Survey Methods in Social Investigation*, Second Edition, C. Moser and G. Kalton, Basic Books, Inc., New York, NY, 1972.
- *Survey Research Practices*, G. Hoinville, R. Jowell and associates; Heinmann Educational Books, London, England, 1978.
- *Survey Sampling: A Non-Mathematical Guide*, A. Satin and W. Shastry, Statistics Canada, 1983.
- *Surveys by Telephone*, R. M. Groves and R. L. Kahn, Academic Press, Inc., New York, NY, 1976.
- *The Art of Asking Questions*, S. Payne, Princeton University Press, Princeton, NJ, 1951.
- *The Dynamics of Interviewing: Theory, Technique and Cases*, R. L. Kahn and C. F. Cannell, John Wiley & Sons, New York, NY, 1957.
- *The Sample Survey: Theory and Practice*, D. P. Warwick and C. A. Lininger, McGraw-Hill, New York, NY, 1975.
- *Understanding Robust and Exploratory Analysis*, D. Hoaglin et al, Wiley, New York, NY, 1983.

☆U.S. GOVERNMENT PRINTING OFFICE: 1985-461-221/24025